INDUSTRIAL R&D AND PRODUCTIVITY IN CANADA

by Frank Longo

A manuscript report prepared
for the Science Council of Canada

May 1984

Manuscript reports are made available on a limited basis to provide information and to elicit comment. They have not undergone the review or editorial process.

85-2984/

Enquiries may be addressed to
the author or to

Science Council of Canada
100 Metcalfe Street
Ottawa, Ontario
K1P 5M1

This paper represents the opinions of the author rather than those of the Science Council of Canada.

ABSTRACT

It has often been argued, particularly recently, that too little industrial R&D is performed in Canada. Consequently the federal government has introduced various policies and programs since the late 1950s, to encourage increased performance of R&D by private firms. In the last five years especially, tax and grant incentives have been sweetened considerably. Other measures, such as weakening competition policy and increased contracting of government R&D to the private sector, have been suggested or tried. This study evaluates the rationale and strategic implications for government support, the factors that affect R&D expenditure and the effects of R&D on productivity.

The study is divided into two sections. The first section examines the factors that determine optimal R&D expenditure. It is found that R&D is bound to be underfunded by the private sector. This is because the market advantage R&D creates for the firm will never generate revenue that is as great as the benefits that flow from the R&D. As well, the factors that cause the underfunding of R&D vary in importance among industrial activities. This suggests that corrective subsidization needs to be selective in order to maximize the benefits.

The second section contains a statistical analysis of the determinants of R&D. This uses a firm-level database developed in cooperation with Statistics Canada. Evidence is found that Canadian-owned firms respond positively to government R&D grants by increasing their own private expenditures on R&D. On the other hand, there is no evidence that weakening competition policy or increasing the contracting out of government R&D will have the same effect. There is also strong evidence that, in addition to not responding positively to government grants, foreign-owned branch plant firms perform significantly less R&D than Canadian-owned firms, other things being equal. In contrast, a separate class of foreign-owned firms whose operations are technologically independent of their parents, perform as much R&D as similarly placed Canadian firms.

This analysis also reveals that R&D and purchased technology are generally complementary. This implies that the maintenance of domestic R&D capability is necessary to get the maximum benefits from the worldwide pool of technology.

In the final part of the study, the effects of R&D on productivity are explored. This is done by estimating the relationship between productivity and the important factors that affect it, including R&D. This reveals that foreign-owned subsidiaries are not more productive than Canadian firms when the important determinants of productivity, such as the capital/labour ratio and capacity utilization rate, are accounted for.

Finally, it was found that the rate of return on R&D in technology-intensive industries is about 24 per cent. This rate is substantially above that for other investments. The estimate is conservative since it does not include the benefits of R&D that occur outside the performing firm. Nonetheless, even this conservative estimate indicates that industrial R&D should be increased by about two-thirds over 1979 levels.

ACKNOWLEDGEMENTS

The author would like to thank the following people for their assistance. Bill Forward worked to have the project initiated and provided important guidance and constructive criticism along the way. Guy Steed patiently oversaw the completion of the project and his insight helped focus its objectives. Paul Gorecki and John Baldwin provided advice on data sources and the direction of the research. John McVey and the staff of the MNEs section at Statistics Canada, especially Barry Mersereau, are much appreciated for their data development activity and their assistance with the computer work. The cooperation of Humphrey Stead and Michel Boucher of the Science Statistics section of Statistics Canada was also necessary. The project benefited from the word processing done by Cher Daley and the important research assistance of Frances Bonney and the Science Council's library staff. Finally, Frank Denton and Tim Hazeldine made critical reviews of early drafts that vastly improved the final result.

The author retains full responsibility for any errors that remain.

CONTENTS

Introduction 9

Section I: An Analytical Framework 11

- Introduction 11
- R&D as a Public Good 11
- R&D and Market Power 13
- Market Structure and Long-Term versus Short-Term Effiriency 16
- Process (Cost-Reducing) Innovations 20
- The Effects of Uncertainty 22
- Summary of the Analytical Framework 23

Section II: Quantitative Analysis 24

- Introduction 24
- The Data 24
- The Determinants of R&D 24
- Firm Size 25
- Market Power 25
- Technological and Financial Opportunity 27
- Foreign Ownership 28
- The Effect of R&D on a Firm's Own Productivity 36

Summary and Conclusions 42

Appendix 1 44

- The Minimum R&D Price 44
- The Sylos Postulate 44
- Derivation of the Limit Price for Homogeneous Products 45
- Market Growth 45
- The Limit Price Under Conditions of Product Heterogeneity 46

Appendix 2 47

- Data Sources 47
- Lag Structure and Depreciation of the R&D Stock 48
- Rate of Return Calculations 48
- An Alternative Estimate of the Production Function 49
- Recommendations for Data Development and Further Research 50

Notes 52

INTRODUCTION

One of the most important issues in industrial policy is the nature of the links between R&D, productivity and market structure. Most analysts agree that market forces left to themselves cannot generate a sufficient amount of R&D and that therefore some public subsidization of R&D is called for. However, there is no agreement on the level of funding required, or the form that assistance should take.

Although the debate remains unresolved, considerable funding of industrial R&D already occurs in Canada and abroad.[1] Canada has been an innovator in the use of tax expenditures to subsidize R&D and it has been found that our tax system is one of the most generous in the world in this respect.[2] With government contracting and grants taken into account as well, Canada's formal R&D subsidization level remains close to the top.[3] In addition, an April 1983 budget paper proposed changes to the tax system (since adopted) that will have the net effect of making R&D investment considerably more attractive.[4]

It is possible to observe most of the relationships that need to be measured in order to make useful R&D policy. However, economists have not yet produced numerical values that they can be confident of within a narrow range for most of the factors that are important in policy-making. This is mainly because of a lack of reliable data and the complicated nature of the interaction between the variables.

Despite these difficulties, certain general hypotheses have been established. The marginal rate of return to R&D has been found, in a number of US studies, to be considerably higher than that for other investments. This is true even when only internal benefits to a firm are accounted for.[5] Estimates of the external effects indicate that these too are very high.[6] As well, R&D expenditures have been found to rise more rapidly than firm size up to a certain point, after which the relationship is at most proportional.[7] It has also been established, in Canada, that foreign ownership has a negative effect on R&D.[8]

Econometric analysis has not yet produced definite answers to a host of more refined, but very important questions. These include the relationship between market power and R&D, the efficacy of government grants for R&D and the extent to which R&D and purchased technology are complementary or substitutable. Finding answers to these questions is not simply an empirical matter. The relationship between R&D and market power especially, remains a controversial issue at the theoretical level.

The objectives of this study are twofold. First, the complexity of the economic interactions between R&D and market power make a deductive analysis necessary to specify the issues and possible policy responses. Secondly, the effects of various determinants of R&D and the effect of R&D on productivity will be estimated.

The paper is divided into two sections. In the first section a theoretical model of industrial R&D is worked out. The reasons for market failure with regard to R&D are developed in the context of a limit-pricing model in which R&D creates an entry barrier. This allows an analysis of the interaction between R&D and market power. The resulting short-term versus long-term trade-off in efficiency and the implications for government policy are discussed.

The second section details an empirical estimate of the determinants of R&D and the effects of R&D on productivity. This was done with a firm-level database, specially constructed for this purpose at Statistics Canada.

The estimate of the rate of return on R&D is the first such attempt in Canada using firm-level statistical analysis. The rate of return on R&D to the performing firm itself is a number with direct policy relevance. Even excluding external effects, R&D is likely to be underfunded and most Canadian R&D is of the cost-reducing type that has its main effect on the productivity of the performing firm. Estimates of the spillover effects of R&D, to firms or industries other than the performer, are not

attempted here. The first Canadian attempt to estimate these external effects has recently been published by the Economic Council of Canada.[9] This was done with an input-output matrix, which is the conceptually correct methodology. Any substantial improvement on this analysis would require the construction of a technology flows matrix (such as the one used by Scherer in the US)[10] to replace the standard Statistics Canada matrix, which is based on production flows. This would probably be a useful project, but a very expensive one.

In this paper the verbal exposition is supplemented by diagrams and equations, but to keep the text as simple as possible, appendices are used to segregate much of the technical detail.

SECTION I: AN ANALYTICAL FRAMEWORK

Introduction

In recent years, the economics of Research and Development (R&D) have received increasing attention. Explicit recognition of the fact that the rate of technological progress is affected by the decisions of economic agents has become common. The replacement of the simplifying classical assumption of an exogenously determined technology level, with an endogenously determined stock of knowledge, necessarily changes the implications for economic policy.

Although knowledge can be treated as just another variable in theoretical models of a firm, market imperfections such as externalities and market power are especially likely to flow from knowledge-creating activity. Hence, the traditional simplifying assumptions become all the more unrealistic when R&D is the subject of analysis.

There are two distinct (though not entirely separable) classes of possible market failure related to R&D. The first of these includes the nature of R&D as a public good, differentials in the variance of return (risk) of other factor investments, and differences in social and private discount rates. The theoretical consequences of these types of imperfections are well established and pretty much undisputed.

The second class of complications involves the relationship between R&D and market power. Product differentiation and economies of scale may produce non-pricetaking behaviour by firms and hence a radically different result from that of a competitive market. This is true even if the remainder of the standard set of neoclassical assumptions hold. R&D can create market power by product differentiation or by changing cost functions, while market power can affect the marginal-revenue-product of R&D. As will be shown below, these factors are particularly important under conditions of uncertainty.

Unlike the first class of complications, there is no consensus on the consequences of oligopoly. This is a contentious matter and a central issue in the literature of industrial organization. Thus it is essential to assess the sensitivity of arguments concerning the optimal level of R&D to various views of the behaviour of oligopolists. In general, there is a welfare trade-off between competition and R&D. Greater price competition lowers consumer prices and improves the allocation of resources. However, it is also held by many to reduce the incentive to innovate. This theoretical trade-off is what is generally known as the "Schumpeterian hypothesis."

In a small economy like Canada's, there is the additional problem of exogenous foreign influences to consider. For instance, it is often suggested that foreign competition can reduce the oligopoly power of domestic firms. Alternatively, the encouragement of domestic R&D could be a way to offset the oligopoly power of foreign exporters to a small domestic market.

R&D as a Public Good

R&D is different from other forms of input into the production process in that it is a public good. Once knowledge has been created, the social cost of its additional use is zero. There are dissemination costs, but the social value of knowledge is not worn out with its use. Thus, for the maximization of social benefits, all useful, non-redundant applications of existing knowledge must be allowed.

This causes market failure, because any price above the dissemination cost, charged for the use of the knowledge, will prevent some worthwhile applications. The higher the price charged, the greater the value of the lost applications. However, if the producers of knowledge cannot charge for it, they will have no incentive to produce it.

Some types of knowledge are considered to be <u>pure</u> public goods. These are goods that have zero marginal cost, and that are not amenable to sale. It is generally agreed that much valuable knowledge is of this nature and should therefore be produced and distributed freely by universities and government or other non-market institutions.

There is, however, much knowledge from which non-payers can be excluded so that it can be marketed effectively and thus characterized as a semi-public or "club" good. The output of industrial R&D is assumed in what follows to be a club good having zero marginal cost and at least temporary excludability.

It is obvious that a firm with useful knowledge or information that can be guarded can earn economic rents on the final product produced, as long as the advantage lasts. However, even if there is sufficient competition to keep the present value of the rents down to the cost of the R&D, there will still be a deadweight welfare loss for society.

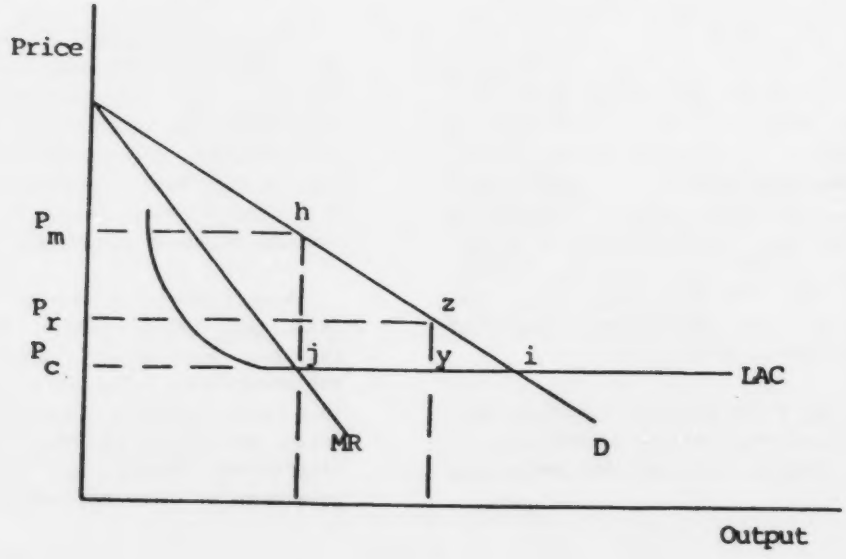
This can be illustrated in a simple model of the firm. Figure 1 shows the case of a firm that makes a new invention as a result of expenditure on R&D. The demand curve and corresponding marginal revenue schedule are labelled D and MR respectively, while LAC is the long-run average cost curve, which includes opportunity cost but is exclusive of the cost of R&D. For simplicity, the LAC curve has been stylized as becoming flat at some discrete minimum efficient scale, instead of asymptotically.

With an effective patent, the firm will have a monopoly and can charge a price P_m, in which marginal revenue equals long-run marginal cost (which is the same as the LAC schedule in the efficient scale range). At this price the firm will earn monopoly profits of $P_m hj P_c$ per period. The net welfare loss per period will be given by the triangle hij. If patent protection is a perfectly effective barrier to competition the firm can earn these monopoly profits until the end of the patent term. It will be profitable for the firm to do the R&D and achieve this position if the present value of the monopoly rents is greater than the cost of the R&D.[11] This means that even an entrepreneur with a monopoly of infinite duration undervalues R&D, since the maximum obtainable profits from R&D are less than the potential welfare gain from the innovation, given by the area under the demand curve and above the cost line.

The patent can thus make R&D possible and allow society to capture some, but not all, of the potential benefits of the new product.

An alternative and more likely assumption is that patent protection is only partial.

Figure 1



Although the patent will stop other firms from producing exactly the same product, a potential competitor can produce a near-substitute with similar expenditure on R&D. If this is the case, then the potential competitor will take some of the demand from the initial firm and earn some of the profits.

It would, of course, be worthwhile for a new firm to enter only if the profits it expects to be earning after entry outweigh the cost of R&D. Because there must be some price above LAC at which the present value of industry profits would just equal the cost of R&D, this price P_r, will give the lowest price (highest output) equilibrium possible. At P_r even an efficient producer with 100 per cent of the market cannot make excess profits. Even at this price there is misallocation of resources. The triangle zyi represents the minimum loss that is the consequence of the nature of R&D.

Equation (1) shows P_r in terms of quantities given by the cost and demand functions. The competitive price and output are P_c and X_c respectively, r is the interest rate, RD the necessary investment in R&D and e is the elasticity of demand around X_c. (The derivation of this and the other equations appearing in this paper are left for Appendix 1).

$$(1)\quad P_r = P_c + P_c[X_c - (X_c^2 - (4.X_c.e.RD.r)/P_c)^{1/2}] /2.X_c.e$$

From equation (1) it is apparent that P_r, when real, is always greater than P_c. If the quantity that is put to the exponent 1/2 is negative there is no real value of P_r, which is to say that there is no price at which the firm can earn profits with a present value as great as the cost of the R&D. Thus no firm will want to develop a product unless the condition given by (1a) holds. A rearrangement of (1a) yields (1b).

$$(1a)\quad X_c^2 \geq (4.X_c.e.RD.r)/P_c$$
$$(1b)\quad RD \leq P_c.X_c/4.e.r$$

Therefore, inequality (1b) gives a necessary condition of existence of a product market. Unless (1b) holds, no firm can earn a net of R&D profit in the market created by the development of the product. Notice that (1b) might not hold even if the cost of R&D is less than the area under the demand curve and above the P_c line. This means that the market will fail to provide sufficient incentive for some socially useful R&D projects to be performed at all.

R&D and Market Power

If the initial firm and potential entrants recognize their interdependence and behave as oligopolists, rather than as naive competitors, the equilibrium price will be above P_r. Bain's concept of a "limit price" is useful here. Bain showed that, given partial entry barriers, there will be a limit price above the lowest possible equilibrium price, which is just low enough to prevent further entry, even without legal restrictions.[12] Bain held that if this limit price is below the monopoly price it can make sense for existing firm(s) to charge this price and earn less than maximum short-run economic profits, in order to prevent entry.[13]

In the absence of other barriers to entry such as economies of scale, absolute cost advantages and brand loyalty, R&D by itself will constitute a barrier resulting in a limit price above P_r, at which profits in excess of the R&D cost can be earned without attracting entry. How much above P_r this price is depends on what share of the market an entrant can take at a price high enough to cover all the entrant's costs, including R&D.

Investment in R&D is analytically distinguishable from other forms of capital investment in this context, because it is rarely proportional to the ultimately foreseen output and it does not depreciate with use.

To determine the limit price it is necessary to know how the existing firm(s) behave(s). Limit-pricing theory assumes that two or more firms will act in collusion as if they formed an explicit cartel.[14] In the particular case being examined - that of

a new product - the first firm in the market has a monopoly, so this controversial assumption is not necessary. The questions that must be answered before the limit price can be determined are: what share of the market can the potential entrant expect to get at each price? and how will the initial firm react to entry?

If the products of the existing firm and the potential entrants are homogeneous, the two questions reduce to one. A firm could not capture any fraction of the market at a price higher than that of its competitors. With a homogeneous product, the initial firm would have to keep pace with a competitor in price, in order to sell anything at all.

The long-run competitive equilibrium shares of firms are indeterminate with a horizontal cost curve, although the price and industry output are not.[15] A unique solution is possible if it is known how the initial firm will change its output in response to entry. Following Sylos-Labini (who extended Bain's work), it will be assumed that the first firm holds output constant in the face of entry.[16] The rationale for this restriction and the effects of its relaxation are discussed in Appendix 1.

In Figures 2a and 2b, the limit price is shown as P_L. The LAC curve is drawn as a horizontal line so that the barrier to entry emanating from R&D can be examined initially in isolation from the effect of economies of scale. P_E is the price that would occur if another firm entered and maximized its profits, while the first firm held output at the pre-entry level.

If the innovator initially sets the price at P_L and produces an output of X_L in Figure 2a, then the demand left for an entrant is given by D_2 in Figure 2b. This is because potential entrants assume the initial firm will bid the price down as far as necessary in order to sell X_L, but no farther. In effect the innovator lets a potential entrant have the part of the demand curve to the right of X_L. The best the imitator can do is charge P_E and earn profits given by the blocked out area in Figure 2b.

If P_L is chosen such that the largest possible profits of an imitator are just less than the cost of R&D, no other firm will enter. Thus the innovator has a choice between charging P_L and earning positive net of R&D profits indefinitely, or charging P_M and earning monopoly profits until

Figure 2a

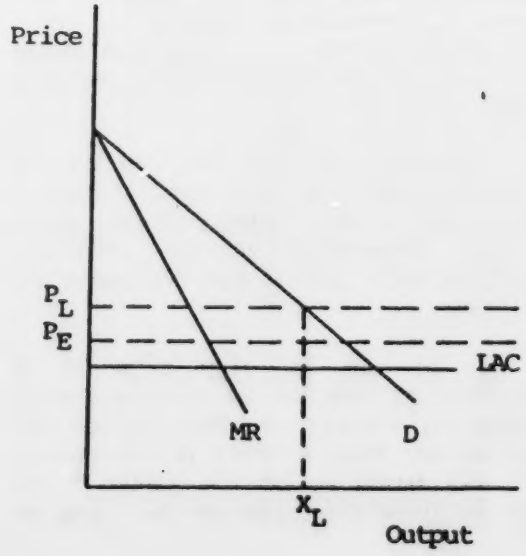
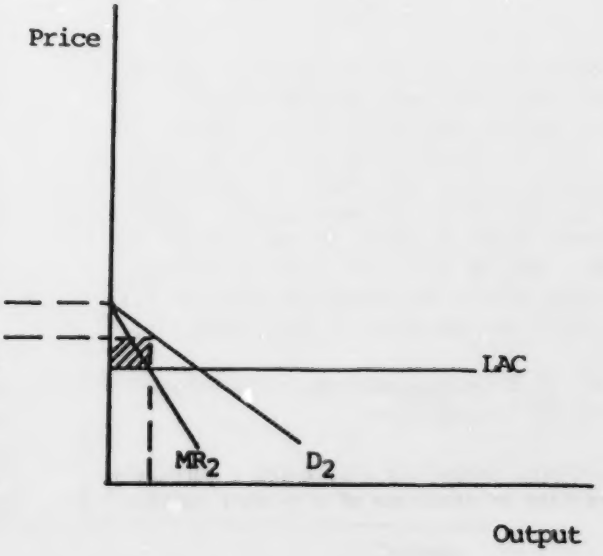


Figure 2b



entry occurs and less than limit price profits thereafter.

If the innovating firm holds output constant, the limit price made possible by the necessity of R&D is given by equation (2) for a linear industry demand curve (see Appendix 1).

$$(2) \quad P_L = P_c + (4a_1 RD.r)^{1/2}$$

As before, P_c is the competitive price, RD is the necessary expenditure on R&D, a_1 is the absolute value of the slope of the demand curve, and r is the interest rate.

Usually, the demand for a new product grows following its introduction. It has been argued that if the market is growing, the price that will prevent entry will be smaller.[17] A growing market can also mean that a firm may create a market even if the profits it expects to make at first are smaller than RD.r.

Consider a case in which the demand curve will grow over time as shown by (3), in which t is time and g is the growth rate.

$$(3) \quad P = (1 + g)^t a_0 - a_1 X$$

If the Sylos postulate is taken literally so that potential entrants expect the innovator to maintain output at its initial state, even after market growth, the limit price will be given by (4).

$$(4) \quad {}_tP_L = P_c + (4a_1 RD.r)^{1/2} + a_0(1 - (1 + g)^t)$$

For positive values of g, no price will keep competitors out indefinitely. Any price below that given by (2) will keep out competition for a finite period of time. The lower the price set, the longer the freedom from competition will be. The innovator will now begin with the no-growth limit price and then continually reduce it just enough to keep competition out. The price at any time will necessarily be lower than that of the no-growth case (see Appendix 1).

Taken in this fashion, the Sylos Postulate becomes an exceptionally optimistic possibility from the point of view of potential competitors, if the growth rate is expected to be large. If the Sylos Postulate is made dynamic an entrant's expectation of the quantity the innovator will put on the market will be given by (5) (see Appendix 1). This specification implies that the innovator will increase production by the same absolute amount that the market grows where market growth is defined by the growth in the competitive equilibrium output.

$$(5) \quad {}_tX_L = {}_oX_L + [a_0(1 + g)^t - a_0]/a_1$$

This version of a dynamic Sylos Postulate yields the same limit price as the no-growth case, given by equation (2). The limit prices of equations (2) and (4) thus represent the results of opposite assumptions about the innovating firm's reaction to market growth.

Although the most reasonable assumption is probably somewhere between these extremes, the dynamic Sylos Postulate has the advantage of consistency over the other extreme. There is no reason to assume that potential entrants take into account information on demand growth and the initial firm does not. This, however, is exactly what is implicit in equation (4).

Implementation lags are another source of variation in the R&D market power relationship. There will probably be a significant lag between the initiation of an R&D project and the production of a saleable good.[18] This will increase the cost of R&D measured at the point at which production begins, which makes it less likely that R&D will occur, because the R&D performer will have to carry the financing charges on the money spent on R&D until the payoff period begins.

As well, the cost of R&D will also rise for potential competitors, which will increase the limit price. Furthermore, the option of charging the monopoly price will become more attractive as the lag length increases, if competitors cannot start imitative activity until the first firm puts its product on the market. This could be the

case if R&D projects can be kept secret until the product is ready for production. Therefore, the longer the lag length, the more unlikely it is that the project will be performed, the higher the price charged and the lower the consequent welfare gain if it is performed.

A firm's ability to maintain its exclusive rights to the knowledge it creates is another important factor in the R&D market power relationship. The social value of R&D does not depreciate with use, although it might be reduced over time if new knowledge makes it irrelevant or redundant. Regardless of how much (if any) social depreciation of an R&D irvestment occurs, the private value of the investment will likely decrease over time as the proprietary knowledge leaks out. This can happen because the employees in whom the knowledge is embodied change jobs or talk to their peers in other firms, or because substitute knowledge development occurs.

This reduces the cost of reproducing the product of R&D as time goes by, which will reduce the entry barrier resulting from the necessity of R&D, as well as the incentive to perform the R&D.

Market Structure and Long-Term Versus Short-Term Efficiency

The importance of the determinants of the limit price for innovation policy stems from the trade-off of long-term versus short-term efficiency. The closer the limit price gets to the monopoly price, the smaller the long-term efficiency loss, but the greater the short-term loss. Conversely, as the limit price approaches the minimum R&D price of equation (1), the current losses are reduced to a minimum, but more socially useful invention is foregone. The ideal situation for society would exist if all R&D projects that create new social surplus greater than their costs were performed, with the resulting knowledge universally distributed so that the competitive price was charged for all products.

However, if such activity is left to the market, this ideal will not be attained. Forcing firms to charge competitive prices for products requiring R&D would give them no incentive to spend money on innovation. As the ability of firms to appropriate the benefits of R&D grows, more innovations with social benefits are performed, but the social benefit of each innovation falls because of static inefficiency. If all products required R&D in the same proportion, the short-term efficiency losses seen in partial equilibrium analysis would be illusory. However, this is certainly not the case. R&D expenditures vary enormously from industry to industry.

The paradox is that, given some monopoly power creation of R&D, the output of R&D-intensive industries will be overpriced while the R&D itself is not adequately rewarded. The obvious answer to this problem is the subsidizing of R&D expenditures. However, the effect of such funding on R&D depends on the firms' ability to retain exclusive use of knowledge. On one extreme is the situation in which firms could keep all knowledge indefinitely. In this case any inventing firm would have a perpetual monopoly. Subsidies would have no direct effect on the price of products that would have been invented anyway. However, previously extra-marginal projects would now be intra-marginal.

If, for example, the demand curves were actually linear, monopoly profits would be half the potential additional social surplus from the invention of a new product. Half this excess social surplus would constitute the normal triangular loss associated with monopoly and would not be recovered through subsidization of R&D since firms would continue to charge the monopoly price. The other half (which is the area under the demand curve and above the monopoly price) could be effectively made internal to the firms profit calculations with subsidization. Without a subsidy, profits would be two-thirds of the obtainable social benefits (assuming the deadweight cost of monopoly cannot be removed). If the subsidy were set at one-third of the R&D expenditures, firms that relate the net cost of R&D to profits would fund a socially optimal amount of R&D.

This efficiency improvement by itself could not be a Pareto Improvement, because some income from taxpayers will be redistributed to firms for previously intra-marginal R&D. This will be true unless projects that were already intra-marginal could be excluded from the subsidization by government economists with extraordinary powers of "cost-benefit analysis."

Consumers would gain only from the additional consumers' surplus of previously extra-marginal R&D, while there would also be a redistribution of income from taxpayers to R&D performers. Assuming that taxpayers and consumers are the same group, their welfare gain would depend on just how much socially useful R&D was added because of the subsidy.

The opposite extreme is the case in which competition is so fierce that it drives the price down to the minimum R&D price given by equation (1). In this case, the price of all goods produced with R&D would fall with subsidization. Consequently there would be a static welfare gain as production of the goods would increase. Since profits are exactly equal to the cost of R&D, the value of the subsidy plus the welfare gain would all accrue to consumers.

In summary, subsidies can result in long-term but not short-term efficiency gains under monopoly, but under competitive conditions they can offer both types of gains. In the in-between case in which an effective limit price exists, some of the static gains obtainable under the highly competitive conditions would be captured.

This analysis has shown that where R&D expenditure is necessary to create a new product, market power will be a concomitant result. It also means that patents that allow a firm to exist as a monopolist will have no effect on the extent of innovation in a certain world. Any innovation that meets the condition of existence will be performed regardless of whether or not it can be patented. The explicit recognition of uncertainty in the model would change this result, since the patent will make the returns to R&D less uncertain and hence more valuable.

The analysis thus far does not, however, allow for market power, other than that created by R&D itself. But an analysis of the special case in which one firm or a consortium has an exclusive ability to make a product is of little use. The relevant case is that in which firms have predetermined and varying abilities to differentiate their products. What is needed is an analysis that allows some partial product differentiation to exist between firms so that policy conclusions for the more likely intermediate case can be reached.

Generally, firms face demand curves that slope downward. If this is admitted, the economy must be regarded not as a group of industries in each of which one or more firms operate, but as a group of monopolists for each of which demand is positively related to the price of imperfect substitutes produced by other firms. Dropping the assumption of homogeneity in the preceding analysis will generalize the model sufficiently to incorporate the "Schumpeterian Hypothesis."

The demand function of the firm inventing the new product can be thought of as a partially reduced form for which the slope coefficient is the net effect of a change in the firm's output on its own price, after accounting for the reactions of other firms' pricing of existing products.

Even when the innovator expects that a potential entrant will produce a substantially different product, the existing firm's choice of price and corresponding output will still affect potential competitors' decision to enter.

This is illustrated in Figure 3 in which the potential entrant is faced with a demand curve that is affected by the price charged by the initial firm. The entrants' price and output axes are labelled P_E and X_E respectively. If the innovator charges price P^1 the demand for the entrants product will be DP^1. At the higher prices

P^2 and P^3, the potential entrants' demand will be higher, DP^2 and DP^3 respectively. With P_c given, some price (P*) will allow an entrant just sufficient demand, DP* with a corresponding marginal revenue schedule MRP*, to make economic profits equal to RD.r. At or above P*, entry will be attracted.

If the shaded box in Figure 3 is equal to RD.r, then P* will be the limit price if potential entrants expect the initial firm to hold the price - rather than the quantity - constant after an imitator enters. If the imitator expects the first firm to hold the quantity constant instead, the limit price will be above P*. This is because the entrance of the imitator will shift the innovator's demand curve down and cause the innovator to charge a lower price after entry, which will in turn shift the entrant's demand curve down.

A simple example, using linear demand curves, will suffice to compare these alternatives. Equations (6) and (6a) show demand curves for the innovator and a hypothetical entrant respectively.

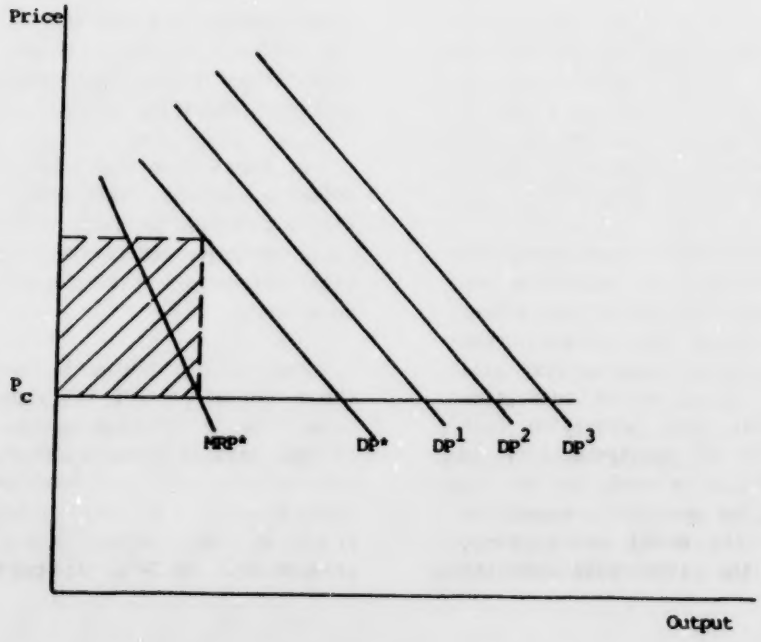
$$(6)\quad P = K^{(1-b_2)}\,(b_0 - b_1X - b_1b_2X_E)$$

$$: K>0 \text{ and } 0 \leq b_2 \leq 1$$

$$(6a)\quad P_E = b_0 - b_1X_E - b_1b_2X$$

In this example P and X are the price and output of the initial firm and P_E and X_E are those of the entrant. Each firm's demand curve is affected by the quantity the other puts on the market. How closely the two demands are related depends on the extent to which their products are homogeneous. Hence b_2 is a parameter of homogeneity that is equal to one when the products are perfect substitutes, to zero when the demands are totally unrelated and to a positive fraction otherwise. Multiplying the first demand curve through by the constant $K^{(1-b_2)}$ allows the usual parameters of the firms' demand curves to be different. The constant K is to the exponent $(1-b_2)$ so the demand curves converge as homogeneity is approached. It can be seen from (6) and (6a) that as b_2 approaches 1 both curves approach the homogeneous case. If b_2 were equal to zero, the equations would describe the demand curves of two independent monopolists.

Figure 3



The limit price when the entrant expects the innovator to hold price constant is given by equation (7), whereas (7a) gives it for an expectation of constant output (see Appendix 1). It can be seen that in (7a) the limit price is greater than in (7) by $K^{(1-b_2)}(RD.r.b_1)^{1/2}$. This is the effect of feedback on the entrant's demand after the price reduction for the innovator caused by entry.

$$(7)\quad P_L = K^{(1-b_2)}/b_2.\ [Pc + 2(RD.r.b_1)^{1/2} - bo(1-b_2) - b_2(RD.r.b_1)^{1/2}]$$

$$(7a)\quad P_L = K^{(1-b_2)}/b_2.\ [Pc + 2(RD.r.b_1)^{1/2} - bo(1-b_2)]$$

As in the homogeneous case, the limit price is positively dependent on the costs of production and R&D faced by the potential entrant, regardless of whether these costs are greater or lower than those faced by the existing firm.

As b_2 approaches one, (7a) approaches the limit price under homogeneity and the Sylos Postulate given by (2). The first derivative of P_L with respect to b_2 is always negative as long as P_L is positive and the condition of existence for entry is met. Thus an increase in product heterogeneity (decrease in b_2) will increase the limit price. As well, the first derivative of P_L with respect to $K^{(1-b_2)}$ is positive. These relationships are the embodiment of the "Schumpeterian hypothesis." As the exogenous component (accumulated distributional advantage) of a firm's demand grows relative to its competitors and its independence of its competitors' increases, the firm's incentive to innovate will increase. That is, its incentive to innovate grows with its market power.

However, this is true only to a point. Although the limit price rises continually with market power, profits rise only to the point at which the limit price and monopoly price are equal. Once the limit price exceeds the monopoly price, the incentive to innovate will not increase with any further increase in market power.

The height of the demand curve faced by different firms for a given product depends on the firms' accumulated investment in distributional activity as well as inherent product quality. Thus a large firm with a strong distributional network could not be kept out of a new market by a small firm that got there first; whereas if nothing but the order of entry were reversed, the large firm could keep the small one out, thus changing the final equilibrium. The fact that the large firm can benefit more from the creation of a given product than the small firm can means that the order of events affects the final equilibrium.

If a small firm carries out an innovation first and larger firms follow it into the market, there will be greater competition and lower prices after the entry process has ended than there would be if a larger firm carried out the innovation first and the number of eventual entrants was lower. Thus the effects of R&D produce a standard infant industry situation. It pays to weight R&D subsidies in favour of newer and less well-established firms, because of the effect on competition.

This kind of argument extends to the situation of a small, partially open economy like Canada's. R&D subsidies should be weighted toward industries facing the greatest problem in competing with foreign producers whose advantage lies in strong market power rather than cost differences.

Another point is that to the extent that domestic producers export, the additional consumers' surplus that results from R&D would accrue to foreign consumers. However, the resulting increase in profits for the firm would be partially acquired from a foreign source. The effect of subsidies on the home country would depend on the market power differential between foreign and domestic firms. Subsidizing R&D can thus create beggar-my-neighbour gains as well as true social gains for the home economy. This could partially offset the subsidizing economy's disregard of the excess social benefits that occur outside its jurisdiction.

Process (Cost-Reducing) Innovations

Thus far we have dealt with the special case of a type of innovation that creates a product that would not have existed otherwise. The cost of production was simply taken as exogenous. However, a large fraction of firms' R&D is designed to reduce the cost of their existing products. It has been found in the US that about 25 per cent of R&D is concerned with reducing the firms' own cost of production.[19] The only large-scale Canadian study that provided an estimate of this fraction found that about 79 per cent of R&D expenditure was spent on process innovations.[20] The difference between these results could be because the Canadian study dealt with only five broad industry groupings, whereas the American data is more comprehensive. However, the higher ratio of process R&D in Canada is consistent with the inherent differences in the two economies. Canada's economy is more oriented toward primary products and has a very large foreign subsidiary component. Both of these factors reduce the incentive for product R&D.

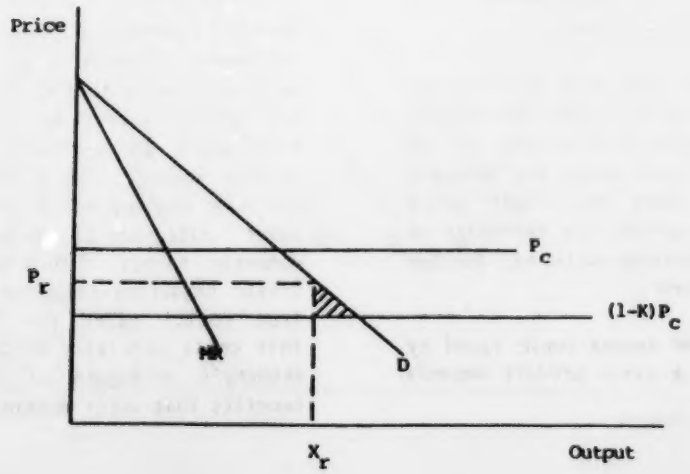
As in the case of product innovation it is pedestrian to show that market failure will arise from process R&D. The cost of process R&D, like that of product R&D, is fixed; there is no marginal user cost. Figure 4 shows the minimum possible deadweight loss incurred by society when firms manufacture a product with a cost curve that is affected by R&D. In the example, an R&D project reduces the cost of production from P_c to $(1 - K)P_c$. Even if the firm charges the minimum price necessary to earn back the cost of R&D (P_r in Figure 4), there will be a deadweight loss given by the shaded triangle. The static loss exists because the entrepreneur must earn back the R&D costs and therefore must charge a price above the competitive price.

Moreover, the profits that will be earned from the R&D, if P_r is charged, will be less than the trapezoid that represents the potential welfare gain made possible by the cost reduction.

What is important of course, is not the potential gain, but rather the actual gain, given that the firms are able to exercise some degree of market power. The actual gain in social welfare, measured by the sum of the additional consumers' and producers' surplus, will always be greater than the increase in profit that accrues to the firm. This is true unless the firm has greater market power than its near competitors.

In the simple partial equilibrium case shown in Figure 5, in which a firm has a downward-sloping and exogenously given demand curve, the excess of the social welfare increase over the increase in

Figure 4



profits resulting from a process innovation, will be equal to $(P_1 - P_2)X_1 + 1/2 (P_1 - P_2)(X_2 - X_1)$. This is the sum of the area between P_1 and P_2, and to the left of X_1 plus the shaded triangle in Figure 5. The triangle represents the new consumers' surplus obtained after the firm's profit-maximizing price is lowered by the cost reduction from P_1 to P_2. The rectangle represents surplus transferred from the producer to the consumer by the price reduction.

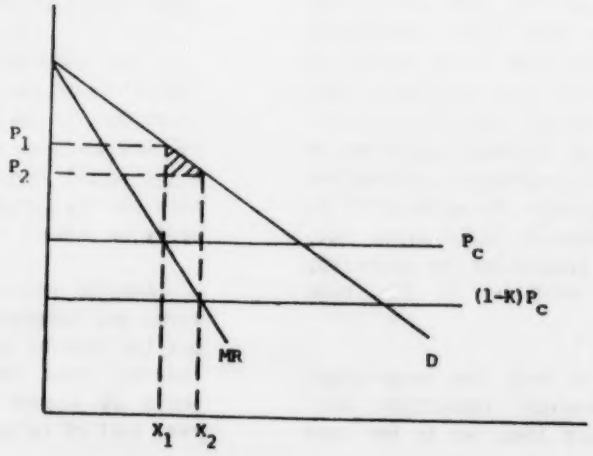
Assuming there are some slightly extra-marginal projects, it will pay to subsidize R&D by the ratio of the additional consumers' surplus to the total increase in welfare. This will reduce the effective cost of R&D to the firm so that the ratio of effective cost to private profits is equal to the ratio of social costs to benefits. A profit-maximizing firm will then perform a socially optimal amount of R&D.

The example above assumes that the firm's competitors do not respond to the price change with lower prices of their own. This could happen if the firm exists in a small, perfectly open economy in which it faces competition only from foreign producers that are very large relative to the home market's size and unable to segment the home market from the world market. In this case, the home country would face horizontal supply curves for the foreign goods, which are not shifted by the domestic producer's activities. The partial equilibrium would then tell the whole story.

It is however, reasonable to assume that the innovator's near competitors can and will respond to the price change. If this happens, there is a negative feedback on the innovator's demand. This effect will reduce the innovators' demand and hence the incentive to innovate. It will also reduce the ratio of private benefits to social benefits resulting from the innovation.

In the general equilibrium, however, there is also another effect that could partially offset the social gains from the cost reduction. This will come about if the firm enjoys more market power than its near competitors. Since the cost reduction will increase expenditures on the product manufactured by the innovating firm, at the expense of competing products, the market power differential between firms will increase, causing a negative welfare effect. Symmetrically, this competition effect will mean that the R&D of firms with little market

Figure 5



power will be more socially underfunded than in the partial equilibrium. This implies that process R&D subsidies should be weighted in favour of smaller, newer firms with less well-established market power.

In discussing the open economy, the general equilibrium is actually more realistic, because there are partial barriers to the flow of goods and services in both directions between any two economies. Bi-directional barriers, whether real or artificial, permit an exporter to charge a different price in the recipient country from the price it charges at home.[21] In this situation, the foreign firm will have a downward-sloping demand curve in the recipient country.

In general, it will be in the exporter's interest to set a separate price in the importing country and to react to the price changes of domestic producers. The foreign producer's price would therefore fall after the domestic producer introduces the innovation. This will mean a redistribution of profits from the foreign firm to domestic consumers, in addition to the gain in static allocative efficiency from the reduction in the market power differential between the foreign and domestic producers.[22]

The salient effects on competition of the process innovation mean that the social rate of return to R&D expenditures of this type will be greater than for the partial equilibrium. Process R&D will be most undervalued in those industries that face competition from foreign producers with a high degree of market power. As in the case of product R&D, some of the excess social value over corporate profit will occur in other countries if part of the domestic producer's production is exported. Nonetheless, it will still be the case that the market undervalues R&D. Unless all of the production is exported, some subsidization will be in the home country's interest.

A final point is that the proprietary knowledge of a process innovation will probably leak out over time, as is the case with product innovation. This will of course reduce the private value over time, so that the market will undervalue R&D by an even larger amount.

The Effects of Uncertainty

Firms do not, of course, actually know the demand and cost curves they face with certainty. Uncertainty makes firms' decision-making less efficient, but this will not systematically affect a firm's decisions if it is an expected value maximizer (risk-neutral). However, it is generally assumed that firms are risk-averse and will therefore discount the value of uncertain profits. Since future events are less certain than more current ones, the result of risk-averse behaviour will be greater discounting of future profits. Thus the more cautious firms are, the higher the price they will charge, since a higher price is reflected in higher current and lower future profits. Furthermore, a higher discount rate will make more projects fail to meet the condition of existence entirely.

It is generally felt that R&D investment is less certain than other kinds of investment because of its nature and its relatively long-term payoff.[23] If firms behave as risk-averse organizations that discount uncertain returns, there will be greater underinvestment in R&D. This conclusion assumes that benefits to society are maximized when non-risk-adjusted returns are maximized, since the risk cancels out over a large enough number of projects.[24]

A mitigating factor in this argument is that firms in the technology-intensive industries that do most of the product-development R&D may be less cautious than the average, which would reduce the misallocation from this source. However, this offset would not apply to process R&D in mature industries.

Another less convincing criticism is that firms may internalize the gains of spreading out the risk by growing larger. This implies however, that the non-social cost of risk would be traded off against the socially real cost of reduced competitiveness.

In general then, uncertainty means that R&D is more socially underfunded than it would otherwise be. As well, the effect of large size on risk reduction means that larger firms should have a higher propensity to spend on R&D. The latter hypothesis is also often associated with Schumpeter. However, there is some opinion that it is more correctly attributed to J.K. Galbraith.[25] For ease of exposition it will be henceforth referred to as the "Galbraithian hypothesis."

Summary of the Analytical Framework

There are several important implications illustrated by this analysis. The hypothesis that R&D will be socially underfunded under any market structure regime is seen to be defensible. When each of the phenomena of leaks of knowledge, lags in development and uncertainty over parameter values are taken into account, social underfunding is all the more likely.

The existence of useful R&D necessarily implies non-pricetaking behaviour by profit-maximizing firms, so that R&D can be said to create market power. This effect is the result of the nature of R&D as a semi-public good. Alternatively, R&D can be affected by market power. These effects create a trade-off between long-term and short-term efficiency objectives.

The social undervaluation of R&D means that it can be socially preferable to subsidize it. Social undervaluation is determined by the extent of uncertainty about the success of new products or processes, the efficacy of patent protection or other impediments to the information flow and the market power of the firm or industry. Since the factors that contribute to the social undervaluation of R&D will probably vary among industrial activities, the optimal rate of subsidization will do so as well.

SECTION II: QUANTITATIVE ANALYSIS

Introduction

There are two empirical questions related to R&D that are of fundamental importance for policy-making purposes: (1) how large would the returns to society be if an extra dollar was spent by firms on R&D? and (2) how do firms change their R&D budgets in response to government assistance or other changes in their decision-making environment? Theoretically, efficiency optimization requires only the equalizing of the marginal social rate of return to R&D and other investments. However, since the amount of privately extra-marginal but socially intra-marginal investment is not known, it is necessary to determine the elasticity of R&D expenditure with respect to changes in its stimuli. If equalizing the social and private returns yields only a small additional amount of expenditure, efficiency will be improved, but social welfare may not be. This is because subsidies entail some redistribution of income from consumers to firms, which establishes a trade-off between efficiency and distributional objectives.

An ideal estimation procedure would involve the use of a completely general model of firm behaviour to estimate simultaneously all channels of interrelationship between R&D and economic performance. The necessary techniques for a simultaneous approach exist (e.g., 3SLS and FIML) and can be easily applied with available software. However, there is a huge gap between the necessary data for this procedure and what is actually available on a comprehensive basis at the level of the firm.

In light of these limitations, a procedure consisting of the estimation of the two policy-relevant reduced forms of this system was followed. R&D expenditure was estimated as a function of exogenous variables, and productivity was estimated as a function of production inputs, including R&D, and other exogenous influences.

The Data

The database used consisted of information on R&D-performing firms in the 24 Canadian, three-digit*, manufacturing industries that had R&D-to-sales ratios of 0.5 per cent or greater in 1979. The concept of the firm used was that of an "unconsolidated enterprise," defined as the aggregation of all the manufacturing establishments in a particular three-digit industry that are commonly owned. All the Canadian numerical data came from Statistics Canada sources. Industry-level data on R&D intensities in the US was obtained from the National Science Foundation. The details of the construction of the database are covered in Appendix 2.

The database includes 447 observations, on which there is data for various periods between 1972 and 1980, depending on the variable. The regressions with R&D as the dependent variable were done for 1979, while the productivity regressions are for 1980.

The Determinants of R&D

Concerning the estimate of this type of equation, a large literature (including some that is specifically Canadian) has emerged. With the cooperation of Statistics Canada, it was possible to construct a more comprehensive and current (1979) data set than was previously available for this purpose. Estimation of the relation between R&D and certain variables, such as government contracts and technology payments, at the level of the firm, was made possible (to the best of the author's knowledge) for the first time in Canada.

There are four factors that have been considered important in the literature for determining a firm's expenditure on R&D. These are firm size, market power, technological and financial opportunity and (especially in Canada) foreign ownership. In this section, the evidence on each of these factors will be examined, in conjunction with this study's approach to the questions raised.

---

* Based on the classification system of Statistics Canada.

Firm Size

There are two reasons to believe that a firm's R&D expenditure will grow more rapidly than its size even if its market power remains constant. These are economies of scale in the performance of R&D itself, and the reduction of risk through multiple investments, made possible by greater size. Empirical evidence suggests, however, that economies of scale in R&D are unlikely to be a factor beyond some fairly small lab size.[26]

Evidence on the relationship between firm size and R&D is mixed. The estimated relation seems to be highly variable with respect to model specification. This indicates that the inability to account properly for factors such as market power that can be correlated with size produces misleading results. However, it is clear that R&D intensity (the ratio of R&D to output) increases with size up to a point (which varies with the industry), but then levels off or declines slightly.[27] The Canadian evidence on this point yields a similar conclusion.[28] If R&D intensity does begin to decline at some point, there must be some anti-synergistic factor at work.

The effect of risk reduction or economies of scale in R&D alone, would be to make R&D intensity increase in size at a decreasing rate. There is thus, at present, some conflict between the theory and evidence on this point.

To get a better estimate of the relationship between R&D and size, data on the manufacturing activity of the firm was used, instead of a figure for total activity. This is because R&D is not necessary for wholesaling operations. Total activity, which includes wholesaling, will probably be an especially poor proxy for manufacturing activity in a small economy like Canada's. In such an economy many firms that manufacture also do substantial wholesaling of imported goods. Therefore, the measure of size used will be TVA, which is the value added for the manufacturing operations of the firm. This and all other variables used in the estimation are defined in Table 1.

Market Power

Attempts to test the "Schumpeterian Hypothesis" have produced little consensus. The empirical analysis has generally consisted of relating innovative activity to measures of the concentration of an industry while attempting to hold technological opportunity constant.[29] Some studies have found a significant positive coefficient to this relationship, but usually with the caveat that the strength of the relationship varies with the technological opportunity of the industry.[30] The reason for the continued uncertainty over this point is undoubtedly that the relationship between market power and innovation is too complex to be easily described.

Recent Canadian evidence on this point is just as mixed. Caves et al. found a strong positive relationship between R&D and concentration at the three-digit industry level for Canada.[31] Howe and McFetridge found no significant relationship between concentration and R&D, in two of three industries studied, using corporate level data.[32] In a study done for the Economic Council with an international sample of two-digit sectors, Palda and Pazderka found mixed coefficients on their concentration variable.[33] The author suspects that this is mainly because of the poorness of the proxies for market power. Industries are, of necessity, defined arbitrarily. Even disaggregated classifications may encompass firms whose products are not very near substitutes. A firm with no near competitors will have a high degree of market power regardless of how many other firms are classified in its industry. Thus measures of industry concentration can give a very misleading picture of market power.

In order to control for market power as well as possible, four variables that give an indication of the market structure of the firm's three-digit industry were included. These are EXPORT1, EXPORT2, PROTECT and HERF, all of which are defined in Table 1. As well, the market share of the firm itself (SHARE) was included as a proxy for market power at the level of the firm.

Table 1

Variable Definitions

FO1	dummy variable, equal to one if the firm is foreign-owned and not classified as an independent cperation, otherwise zero
FO2	dummy variable equal to one if firm is classified as specialized mission, otherwise zero
TVA	total value added of manufacturing establishments in unconsolidated enterprise
PROFIT	gross profits of unconsolidated enterprises manufacturing operations
GRANTS	total of government R&D grants allocated to unconsolidated enterprise
CONTRACTS	total of government R&D contracts
TECHPAY	total technology payments
TECHREC	total technology receipts
EXPORT1	exports divided by sales for three-digit industry
EXPORT2	exports divided by exports plus imports, for three-digit industry
PROTECT	effective rate of protection for three-digit industry
CONSDUR	dummy variable equal to one if industry is considered a producer of consumer durables
CONSNDUR	dummy variable equal to one if industry is considered a producer of consumer non-durables
HERF	Herfindahl index for three-digit industry
TRD	total current and capital and extramural R&D allocated to the unconsolidated enterprise
PTRD	TRD minus GRANTS minus CONTRACTS
ASSETS	total ASSETS of corporate unconsolidated enterprise
PHW	person-hours of labour used by the unconsolidated enterprise
MATERIALS	total cost of materials for the unconsolidated enterprise, inclusive of energy costs
PRODUCTION	value of production, equal to TVA plus materials
UTILIZE	proxy for capacity utilization, equal to the real value of 1980 production, divided by the highest level of real production between 1975 and 1980
SECTOR (1-9)	dummy variables equal to one if the firm is in the corresponding two-digit sector

Table 1 (continued)

RDSTOCK	the stock of R&D, computed as the depreciated total of expenditures between 1972 and 1979
SHARE	output of the firm divided by the total output for its three-digit industry
USRD	R&D-to-sales ratio of corresponding US three-digit industry

Technological and Financial Opportunity

Holding technological opportunity constant is probably the most vexing problem associated with estimating the determinants of R&D expenditure. This is because it is difficult to find a good proxy for the largely unobservable factors that represent technological opportunity. If technological opportunity is not accounted for, the estimated parameter values of other variables may suffer from misspecification bias, because of multicollinearity.

In attempting to hold technological opportunity constant, some researchers have developed indices based on subjective evaluations of an industry's basic science endowment or its distribution of scientists and engineers.[34] Another method is to use industry dummy variables to pick up industry differences in technological opportunity. All of these methods have their drawbacks. For instance, an index based on a factor analysis of scientists and engineers would suffer from simultaneous determination and indeed partial identity with the dependent variable of R&D. Industry dummy variables, on the other hand, will probably pick up spurious differences having nothing to do with technological opportunity. Another approach consists of confining the analysis to one industry at a time, on the assumption that within an industry technological opportunity is fairly constant. For this assumption to be plausible, the industry would have to be narrowly defined, which severely limits the number of observations. Even at the three-digit level (at which a large amount of agglomeration remains) most industries do not include nearly enough observations for a useful regression analysis. The number of observations could be increased through the pooling of cross-sections from different years, but this requires the unlikely assumption that technological opportunity within an industry remains constant over time. This latter assumption is particularly difficult to rationalize because of its inconsistency with any kind of product life-cycle hypothesis.

Yet another approach is to use the R&D-to-sales ratios of corresponding American industries as a proxy for technological opportunity. This proxy, first introduced by Caves et al., is exogenously determined and non-subjective.[35] Hence, USRD (the R&D-to-sales ratio of the corresponding US industry) is included in the regressions as a proxy for technological opportunity. This variable is generally at the three-digit level, but the two-digit ratio is used where three-digit figures could not be obtained because of US confidentiality restrictions.

The analysis of the previous section shows that the incentive for R&D expenditure should increase with the potential for product differentiation. This result, which has appeared repeatedly in the previous literature, has led to the suggestion that consumer oriented and durable goods industries should display higher R&D intensities, other things being equal.[36] Therefore, CONSNDUR and CONSDUR are included. These are dummy variables equal to one if the three-digit industry in which the firm operates is considered to be primarily a producer of non-durable consumer goods or durable consumer goods, respectively. Producer goods are thus the base case for the product differentiation vector.

Another important factor in a firm's determination of optimal R&D expenditure is the availability of outside technology to be purchased. This is likely to be a particularly important factor in a small, relatively open economy like Canada's. The more cheaply outside technology can be purchased, the greater the opportunity cost of R&D expenditure becomes. Alternatively, R&D can make outside technology more easily absorbable. As well, sellers of technology often accept payment in kind so that innovation could make outside technology more available.[37]

Thus TECHPAY (payments outside the firm for technology) is included in the regression. The null hypothesis to be tested by the sign and significance of this variable is that R&D and purchased technology are neither net complements nor substitutes. A positive sign would indicate net complementarity and a negative sign, substitutability.

As well as the availability of outside technology, the extent of the firm's opportunity to sell its technology is a factor in its R&D decision. The variable TECHREC (receipts of the firm for technology) is included as an indication of this kind of opportunity.

The variables used to measure financial opportunity are PROFITS, CONTRACTS and GRANTS. Some previous researchers have contended that because there is two-way causation between a firm's R&D and its profits, a simultaneous systems technique is required to estimate the determinants of R&D.[38] Certainly, in addition to the influence of current and past profits on R&D, R&D expenditures can affect future profits. However, bi-directional causation is not simultaneity. Unless breathtakingly short lags between R&D activity and its effects on products or productivity are common (and there is considerable evidence that they are not), current R&D will not affect current profits.[3] Therefore profits are considered exogenous in this relationship.

The dependent variable in these regressions is PTRD: the current year's expenditure of the firm on R&D of all types, minus the value of government grants and contracts received by the firm. Hence, a negative (positive) coefficient on either the GRANTS or CONTRACTS variable would indicate that firms reduce (increase) their privately financed expenditure on R&D in response to that type of government support.

Several Canadian studies have measured the effect of grants on R&D. All have found a significant positive relationship for at least some portion of the industries studied.[40] However, only one of these, Howe and McFetridge, interpreted this as a ceteris-paribus effect.[41] The others rejected the interpretation on the grounds that other elements of technological opportunity could not be properly accounted for. The inclusion of USRD and other variables which are indicative of technological opportunity should substantially reduce this possible mis-specification bias.

Foreign Ownership

The hypothesis that a foreign-owned branch plant operation will probably do less R&D than a similarly placed domestic firm, is now widely accepted.[42] This is expected because a subsidiary of a foreign-owned firm that produces a subset of its parent's product line can draw on its parent's know-how. The recent Canadian evidence bears out this hypothesis.[43]

The lower propensity of foreign-owned firms to innovate depends on the extent to which they are replicas or components of their parents. This suggests that a foreign-owned firm that is not of this type, but rather an essentially independent operation, should spend the same amount on R&D as a domestically-owned firm.

What is important here is not the managerial independence of the firm, but rather the extent to which their line of business is technologically different from their parent's. If the subsidiary's operations consist of the same or part of the same production process that the parent uses, there will be little reason for the subsidi-

ary to perform R&D other than for adaptation to the local market. Thus the purely technological explanation for the observed behavioural difference implies that the effect will disappear if the subsidiary's operations are widely removed from its parent's.

A survey of foreign-owned manufacturers done by the Ontario Ministry of Industry and Tourism (now the Ministry of Industry and Trade) produced a list of firms that described their operations as largely independent of their parents. The usual phrases for this description were "world (or North American) product mandate" and "specialized mission." This list was used to identify a separate class of foreign-owned firms in order to test the above hypothesis. As a result there are two dummy variables of foreign-ownership used in this analysis. FO1 indicates a firm that is foreign-owned and not on the list and FO2 is for firms that are on the list.

The survey included only firms with 50 or more employees that have operations in Ontario. Most foreign-owned manufacturers operating in Canada do at least some business in Ontario so the less than complete coverage is not likely a major problem.

However, the survey was entirely voluntary and its usefulness depends on the extent to which the descriptions of the business managers are consistent. Thus, caution must prevail in interpreting the distinction between the two categories of foreign-owned firms.

The regression specification contains both dollar variables and scale-free variables such as USRD. This can be misleading if the regressions are specified linearly, since a given change in a ratio such as USRD would be specified as causing a constant dollar value change in R&D, regardless of the size of the firm. This problem will not occur if the equation is specified as being log-linear.

There is, however, no theoretical basis for assuming that the relationship between R&D and most of its determinants is non-linear. Thus, the use of a log-linear functional form to accommodate scale-free variables is a departure from the discipline of Occam's Razor. However, even if a variable is linearly related to R&D, the log-linear form is sufficiently flexible to provide a reasonable approximation of the relationship.

Another problem with the use of a log-linear specification is the occurrence of zero or negative values among the observations of the independent variables. Some of these variables such as GRANTS and TECHPAY have large numbers of zero values. However, most of the variables are positive semi-definite, and in fact, only PROFIT and PROTECT have any negative values. There were not many of these and without exception all the negative values were small in absolute magnitude in comparison to the mean values of the variables. PROFIT represents gross profits rather than net profits, which is why it contains few negative values.

The zero and negative values were assigned an arbitrarily low value (.01) for the purpose of taking logs. This undoubtedly will impart some bias to the parameters. The sensitivity of the results to this problem was tested by using a different approximation of zero (.0001). There was no significant difference between the regression results with the different alternatives.

All things considered, the advantages of the log-linear specification seemed to outweigh the disadvantages. Thus, the logs of all the variables (with the exception of the zero/one dummies) have been taken for use in the regressions reported below.

PTRD is the dependent variable in the R&D regressions. This includes all current and capital R&D expenditures for in-house and contract research. It is netted with government R&D grants and R&D contracts to give a figure for privately financed and used R&D.

The frequent concern of data-mining over a large number of variables did not pose a problem in these regressions. This is

Table 2a

Regression for All Firms
with PTRD as the Dependent Variable

Independent Variable	Coefficient	T - value
Intercept	3.52	(1.40)
FO1	-1.32***	(-4.32)
FO2	-0.36	(-0.80)
TVA	1.82***	(5.16)
PROFIT	-0.24**	(-2.02)
EXPORT1	0.17	(0.67)
EXPORT2	-0.31	(-0.72)
PROTECT	0.05	(0.43)
SHARE	-0.74**	(-2.30)
HERF	0.18	(0.74)
GRANTS	0.04	(0.95)
CONTRACTS	-0.09**	(-1.98)
TECHPAY	0.08**	(2.42)
TECHREC	0.11*	(1.85)
CONSDUR	0.38	(0.69)
CONSNDUR	1.41**	(2.26)
USRD	0.24	(1.58)

R^2	0.32
F-value	12.65
Observations	447
Degrees of Freedom	430

* - Significant at the 90 per cent confidence level (two-tailed test)
** - Significant at the 95 per cent confidence level (two-tailed test)
*** - Significant at the 99 per cent confidence level (two-tailed test)

Table 2b

Regression for Canadian Firms
with PTRD as the Dependent Variable

Independent Variable	Coefficient	T - value
Intercept	-1.71	(-0.55)
TVA	1.37***	(3.19)
PROFIT	-0.02	(-0.14)
EXPORT1	0.26	(0.78)
EXPORT2	-0.15	(-0.29)
PROTECT	0.12	(0.81)
SHARE	-0.75*	(-1.94)
HERF	0.54**	(1.97)
GRANTS	0.10**	(2.24)
CONTRACTS	0.05	(0.96)
TECHPAY	0.08	(1.40)
TECHREC	0.15*	(1.77)
CONSDUR	0.09	(0.14)
CONSNDUR	1.55*	(1.93)
USRD	0.20	(1.27)

R^2	0.38
F-value	8.26
Observations	202
Degrees of Freedom	187

* - Significant at the 90 per cent confidence level (two-tailed test)
** - Significant at the 95 per cent confidence level (two-tailed test)
*** - Significant at the 99 per cent confidence level (two-tailed test)

Table 2c

Regressions for Foreign Firms
with PTRD as the Dependent Variable

Independent Variable	All Foreign Firms		FOl Firms	
	Coefficient	T-Value	Coefficient	T-Value
Intercept	6.41*	(1.67)	8.12*	(1.80)
FO2	1.19**	(2.34)	-	-
TVA	1.98***	(3.59)	1.28**	(2.09)
PROFIT	-0.45**	(-2.31)	-0.42*	(-1.77)
EXPORT1	0.03	(0.07)	0.36	(0.82)
EXPORT2	0.27	(0.37)	0.66	(0.82)
PROTECT	-0.01	(-0.06)	0.18	(0.78)
SHARE	-0.50	(-0.99)	0.16	(0.27)
HERF	-0.54	(-1.18)	-0.90*	(-1.75)
GRANTS	-0.02	(-0.38)	-0.11	(-1.15)
CONTRACTS	-0.19**	(-2.36)	-0.07	(0.50)
TECHPAY	0.08*	(1.65)	0.07	(1.21)
TECHREC	0.13	(1.45)	0.22*	(1.82)
CONSDUR	0.82	(0.84)	1.32	(1.20)
CONSNDUR	0.75	(0.74)	1.11	(1.01)
USRD	0.66*	(1.82)	0.62	(1.57)
R^2	0.34		0.33	
F-value	7.93		6.07	
Observations	245		186	
Degrees of Freedom	229		171	

* - Significant at the 90 per cent confidence level (two-tailed test)
** - Significant at the 95 per cent confidence level (two-tailed test)
*** - Significant at the 99 per cent confidence level (two-tailed test)

because there was a large number of observations relative to the number of variables considered, so that it was possible to estimate all coefficients at once. However, the many ways in which the data could be segmented did create a regression strategy conundrum. It was decided to do a regression with all the observations included and separate regressions for Canadian and foreign firms.

Parameter estimates are characterized as significant if the corresponding t-value is statistically significant at the 90 per cent confidence level. This is determined with a two-tailed test. T-values and degrees of freedom are of course provided for the reader who wishes to apply different criteria.

The regression results for the determinants of R&D are shown in Tables 2a to 2c. Four separate regressions are reported. In the first, all firms in the database were included. The second regression was done only on Canadian firms, while the third and fourth are for all foreign firms and foreign firms exclusive of those represented by FO2, respectively. The results (with a couple of exceptions) are generally consistent with those reported in other studies and the predictions of theory. For convenience the parameter estimates will be discussed in roughly the order in which they appear in the tables.

The estimate of the effect of foreign ownership is as predicted. In the all-firms regression, the coefficient on FO1 is negative and highly significant. This indicates that, other things being equal, a foreign-owned firm of the first type (branch plant firm) does perform less R&D than a Canadian firm. Conversely the parameter estimate for FO2 is not significantly different from zero. This indicates that the other type of foreign-owned firms (specialized missions) are not less R&D-intensive than Canadian firms. This result is confirmed by the positive significance of the coefficient on FO2 in the regression for all foreign firms.

Analysis of the differences between the other estimated parameters, as they are allowed to vary among the ownership categories, does not give any general indication that the elasticities between R&D and its determinants are more similar between Canadian firms and specialized missions than they are between Canadian and branch plant firms. For five of the 10 coefficients that attain significance in at least one of the regressions, the difference between the values for Canadian firms and all foreign firms is lower than that between Canadian firms and branch plants. For the other five, the difference between Canadian firms and branch plants is lower. Only two of these differences are very striking. The coefficient on CONTRACTS is negative and significant in regressions that include specialized missions, but insignificant elsewhere. The coefficient on TVA is much higher in the regressions including the specialized missions. These differences will be mentioned again as each parameter estimate is discussed in turn.

In this log-linear specification, a coefficient greater than one on TVA implies that R&D expenditure increases more rapidly than value added. This coefficient is significantly greater than zero in all the regressions, but significantly greater than one only in the regressions that include specialized missions. This indicates that except among specialized missions there is no strong tendency for R&D intensity to increase with size.

The effect of PROFIT is negative and significant in all the regressions except that for Canadian firms alone, where it is insignificant. This seemingly odd result is consistent with that obtained by Caves et al., who found a negative effect of profit on R&D.[44] Howe and McFetridge, on the other hand, found a positive effect of profit on R&D among Canadian firms and none for foreign firms.[45] American studies have usually found no significant effect of profitability on R&D.[46] This also occurred in the Palda and Pazderka study on international data.[47] Caves et al. offered a few plausible explanations for the negative effect that they found.[48] It would appear though, that the uncertainty of the various

results means that nothing definite can be said on this point.

The two proxy variables for comparative advantage, EXPORT1 and EXPORT2 and PROTECT (the effective rate of protection) are insignificant in all the regressions. The results for the trade variables are consistent with those of Caves et al., whose study was the only other that estimated, with Canadian data, the effect of trade intensity on R&D while attempting to hold the other determinants of R&D constant.[49] On the other hand, that study found a positive relationship between the effective rate of protection and R&D.[50] One explanation for this divergence is that this study deals with only the relatively R&D-intensive industries, whereas the earlier study was done with data drawn from across the spectrum of manufacturing industries. Governments generally try to encourage the more technology-intensive "end-product" industries, so the relationship between tariff protection and R&D may be the reverse of that postulated. This could explain the finding that industries enjoying high protection tend to be more technology-intensive, but that within the technology-intensive group there is no significant relationship.

The results for the market structure variables are mixed, but in general do not lend any support to the idea that there is a positive relationship between the incentive for R&D and the concentration of the industry and share of the firm.

In the all-firms regression, the coefficient on SHARE is negative while that for HERF is insignificant. For Canadian firms the effect of SHARE is negative, and that of HERF positive. Conversely, HERF has a significant negative effect in the regression for branch plants.

The tendency in earlier studies was to interpret this sort of finding as direct evidence on the Schumpeterian hypothesis. This interpretation assumes that standard industrial classifications are composed entirely of firms selling near substitutes for each other's products. This is unlikely at the two-digit or three-digit industry level. Firms within an industrial classification often exploit distinct market niches. As a result, concentration and share indicators may be poor proxies for market power. Thus, the results of this and other studies do not refute the Schumpeterian hypothesis.

The effect of GRANTS on private R&D expenditures is positive for Canadian firms, but insignificant for foreign firms. The coefficient is positive and significant in the regression for Canadian firms only, and insignificant in the other regressions. A positive coefficient on GRANTS indicates that firms increase their own privately financed expenditures on R&D in response to the government grants.

The first derivative of PTRD with respect to GRANTS is equal to the log-linear coefficient (the elasticity between the variables) multiplied by the ratio of PTRD to GRANTS. Since the coefficient is 0.10 and the ratio of the variables at the means is 10.76, the calculated first derivative is 1.07 at the means. Thus the estimate implies that Canadian firms will increase their own expenditure on R&D by about a dollar for every dollar of subsidy.

These results - including the finding that foreign firms are less responsive to grants than Canadian firms - offer rough confirmation of those obtained by Howe and McFetridge.[51]

The coefficient on CONTRACTS is negative in the all-firms regression, contrary to expectations. Among Canadian firms and among branch plants there is no significant effect of CONTRACTS; whereas the coefficient is negative and significant in the regression for all foreign firms. This implies that the negative effect of CONTRACTS is limited to specialized missions.

The hypothesis being tested by the presence of CONTRACTS is that government contracting of R&D to private firms will cause them to do more R&D themselves, since through the realization of economies of scale the firms' marginal cost of R&D will

be reduced. Obviously the results imply that this is not the case.

The negative coefficient among specialized missions could be the result of the government purposely directing contracts toward technologically poor firms to help them overcome the hypothetical economies of scale. However, there is no reason why this should be true of one class of firms and not the others. Thus that peculiar result must be considered a statistical anomaly until more evidence on this point is obtained.

The coefficient on TECHPAY is positive and significant in the all-firms and all-foreign-firms regressions. In the other two regressions the variable falls slightly short of significance, but the value of the coefficient is virtually identical across all the regressions. This is fairly strong evidence that R&D and purchased technology tend to be net complements.[52]

The effects of the last four variables have no obvious policy relevance. As explained above these are included in the regression for the purpose of holding technological opportunity as constant as possible. The signs of the coefficients were all expected to be positive and indeed all four variables have positive signs in each of the regressions.

The coefficient on TECHREC is significantly positive in three of the four regressions and almost significant in the fourth. The unsurprising conclusion is that a firm's R&D expenditures are positively related to its success in selling past innovations.

The dummy variables representing the product class of industry, CONSDUR and CONSNDUR, are consistently positive, but generally insignificant. CONSNDUR is significant in the regression for all firms and among Canadian firms specifically, but CONSDUR is not significant in any of the regressions. Altogether, the results on these variables suggest, though weakly, that makers of consumer goods have a greater incentive to perform R&D than those that make producer goods. The difference is most pronounced among the producers of non-durable consumer goods.

This result is as expected, since consumer goods and especially non-durable consumer goods are generally thought to be easier to differentiate than producer goods. The finding thus represents weak confirmation of the theoretical result developed in Section I, that a firm's incentive for R&D rises with its ability to differentiate its product.

Finally, the coefficient on USRD is always positive, as expected, and significantly so in the regression for all foreign firms. It is nearly significant in the other regressions.

Before turning to the effects of R&D on productivity, a short summary of the findings concerning the determinants of R&D should be useful.

In general, foreign-owned firms do less R&D than their Canadian counterparts, other things being equal. However, this is not true for a separate group of foreign firms that describe themselves as specialized missions. R&D expenditures increase at least proportionately with the output of the firm. Among specialized missions, R&D increases more rapidly than output. There is no obvious explanation for the difference between this group of firms and the others on this point.

There is no statistically significant relationship between R&D and the trade variables or the effective rate of protection. There is also no consistent relationship between R&D and industry concentration or market share.

The results indicate that Canadian firms are responsive to government grants but foreign subsidiaries are not. Alternatively there is no positive response to government contracting. Finally the evidence suggests that purchased technology and R&D are net complements.

The Effect of R&D on a Firm's Own Productivity

It has often been suggested that much of the productivity improvement from R&D will probably show up in industries other than the one performing the research. This hypothesis has been successfully tested and found to be correct.[53] The testing involves tracing the product of R&D through an input-output matrix to its final users to produce a figure for "used" as opposed to performed R&D.

With one exception these studies were done with US data. The most recent of these, which had the advantage of a specially constructed "technology flows" matrix, was done by F.M. Scherer.[54] Scherer found that the social rate of return on R&D was about 100 per cent per year.[55] Earlier studies of this type found rates of return of a similarly high level.[56] The Canadian example of this type of study was done recently by the Economic Council of Canada. This also found a significant extra-private effect of R&D on productivity, although this study did not have the advantage of the specialized and disaggregated database available to Scherer.[57]

It is, however, highly desirable to have an estimate of the firm's own rate of return on R&D. This is because theory implies that R&D would be underfunded even on this criteria, owing to the relative riskiness of R&D as an investment and the excess of private over social depreciation. As well, the ability to derive estimates that are useful for policy purposes from the input-output analysis is limited in Canada at this time by the lack of a database with which technology flows can be traced.

Most of the research on the effect of R&D on a firm's own productivity has also been done for the US. In fact no comprehensive firm-level study has been done for Canada. Industry-level regressions consistently produce no significant relationship between R&D and the performing firm's productivity. For instance, S. Globerman, who used two-digit-level data, was unable to find any positive coefficient on R&D until capital was left out of the equation.[58]

This experience with industry-level data could hardly be more misleading, as the American firm-level literature shows.[59] Time and again these studies produced highly significant relationships between R&D and the performing firm's own productivity. The range of estimates of the rate of return on R&D within a firm vary from about 20 to 55 per cent, well above the similar estimates for the rate of return on capital.[60]

In a recent paper, Bernstein and Nadiri proposed that the reason for this gap is that the adjustment costs involved in R&D accumulation are greater than those for physical capital accumulation.[61] Once adjustment costs are accounted for, the long-run marginal rates of return to R&D and physical capital should both be equal to the interest rate. This condition is expected because it is a corollary of long-run competitive equilibrium. Bernstein and Nadiri did find that their estimate of the adjustment costs accounted for a substantial part of the difference between the interest rate and the rate of return to capital.[62] However, their estimates of the adjustment costs did not explain the higher rate of return to R&D. In one of the four industries studied, the estimated adjustment costs for R&D were much higher than those for physical capital, whereas in one other the opposite was true.[63] In the remaining two industries, there was little difference between the adjustment costs associated with the different types of investment. Thus the residual differences between the rates of return for R&D and physical capital were left largely unexplained.

The best explanation for the continuing difference is that risk-adjustment factors differ systematically for investment in R&D and for physical capital. Firms do not equalize expected rates of return on investment alternatives. Instead they equalize risk-adjusted, expected rates of return. If the risk is higher with R&D investment, the marginal rate of return on R&D will be set by firms above that for physical capital and both will be above the interest rate, since all investment carries with it some risk.

The procedure for the estimation is fairly straightforward. Output is specified as a function of capacity utilization, capital, labour, materials and the "stock of knowledge." It is assumed that the stock of knowledge is a function of the firm's accumulated investment in R&D as well as the general technology level of society. Finally dummy variables are used to pick up systematic industry differences at the two-digit level, and any effect of foreign ownership.

The production function is specified as a Cobb-Douglas, following most of the previous literature. In contrast to the situation with the equation for the determinants of R&D, the application of Occam's Razor suggests the use of the Cobb-Douglas form for the production function.

Another well-known and easily applicable specification for the production function is the translog form. This includes the Cobb-Douglas as a nested hypothesis, while allowing more leeway for the data to determine the shape of the isoquants through squared and cross-product terms in the logarithms. This specification was tried, but an ANOVA test indicated that the additional variables did not significantly improve the fit. Thus the assumption of unitary elasticity of substitution, implicit in the Cobb-Douglas specification, seems to be quite reasonable. Given that the regression is done for a cross-section and hence estimates the long-run parameters, this is not surprising.

The form of the Cobb-Douglas function is shown in equation (12); where A is some constant. For estimation, both sides are divided through by PHW (person-hours worked), to make the dependent variable into production per hour. The natural logs are then taken as shown in equation (12a).

(12) $\text{PRODUCTION} = A(\text{ASSETS})^{b_1}(\text{PHW})^{b_2}(\text{RDSTOCK})^{b_3}(\text{UTILIZE})^{b_4}(\text{MATERIALS})^{b_5}$

(12a) $\log_e(\text{PRODUCTION/PHW}) = \log_e(A) + b_1\log_e(\text{ASSETS}) + (b_2-1)\log_e(\text{PHW}) + b_3\log_e(\text{RDSTOCK}) + b_4\log_e(\text{UTILIZE}) + b_5\log_e(\text{MATERIALS})$

The regressions are done for a cross-section of 146 firms at 1980, for which continuous R&D data back to 1972 could be obtained.[64] The stock of R&D (RDSTOCK) was accumulated R&D expenditure from 1972 to 1979 (the range of available data) adjusted for an assumed depreciation rate of 12.5 per cent per year. A number of different lag structures and depreciation rates were used to produce alternative estimates of the R&D stock and then tried in the equation. The best fit was obtained with the one outlined above, though there was no statistically significant difference between the various alternatives tried. The other depreciation rate and lag structure combinations that were tried are shown in Appendix 2.

The production function can also be estimated in a "rate of change" form, which eliminates the necessity of estimating the stock of R&D. Both approaches have advantages and disadvantages. This is discussed at length in Appendix 2.

The flow variables PRODUCTION, PHW and MATERIALS are all taken at 1980. ASSETS is the accounting concept of total assets net of accumulated depreciation, taken from financial statements (see Appendix 2). The values for both 1980 and 1979 were tried in the regression. As was expected, the use of the 1979 value produced a slightly better fit. This is because the current year's additions to the capital stock are unlikely to affect current production. Thus, for all the results reported, the 1979 value of ASSETS was used.

The inclusion of capacity utilization (current output divided by capacity) means that the coefficients on the other variables cannot be interpreted as ceteris-paribus effects. This is because an increase in the amount of input not only raises output directly, but also indirectly, because at the higher output (with capacity given) the whole production process is more efficient. In order to determine these effects, output per hour must first be isolated on the left side of the equation.

The capacity utilization term in equation (12a) is $b_4 \log_e(\text{UTILIZE})$. This can be rewritten as $b_4 \log_e(\text{PRODUCTION}) - b_4 \log_e(\text{capacity})$. To isolate production per hour on the left side we subtract $[b_4 \log_e(\text{PRODUCTION}) - b_4 \log_e(\text{PHW})]$ from both sides. This will make the left side $(1-b_4) \log_e(\text{PRODUCTION/PHW})$. We then divide through by $(1-b_4)$. Thus for the i^{th} input the ceteris-paribus log-linear coefficient is $b_i/(1-b_4)$ which will be greater than b_i, as long as b_4 is positive and less than one.

The regressions are done with UTILIZE included as an explanatory variable, thus in order to determine the rates of return to the factors, the calculations explained above must first be performed on the relevant coefficients. The alternative would be to make the transformation prior to estimation, which would mean that capacity would replace UTILIZE on the right side.

Table 3 shows two different regressions. In the first, the coefficient for R&D is not allowed to vary among industries, while in the second it is allowed to vary, with the use of dummy variables, for each of the five two-digit industries that had relatively large numbers of observations.[65] In both regressions the dummy variables for sectors and foreign ownership are allowed to change the intercept.

Given the significance of the regressions, as indicated by the F-tests and R^2 statistics, it is reasonable to conclude that the determinants of productivity are fairly well captured. Also encouraging is the fact that the coefficients on ASSETS, MATERIALS and PHW sum to almost exactly the value of zero. This indicates near constant returns to scale, which is predicted by theory and consistent with past results.[66] In both regressions all the ordinary inputs are highly significant as would be expected.

The coefficient for foreign ownership is insignificant in both regressions. This indicates that foreign-owned firms are not more productive in general than Canadian firms using the same input combinations.

What is of particular interest is that the stock of R&D is shown to be a significant factor of production in the first regression. The second regression, which allows this coefficient to vary among two-digit sectors, shows that the significance of this relationship comes mainly from the electrical products and chemical products major groups, SECTOR(3) and SECTOR(7), respectively.

Since the coefficient on UTILIZE is approximately 0.22, the adjustment factor for the partial regression coefficients of the inputs is 1 /(1 - 0.22) or 1.28. To produce an estimate of the rate of return on R&D, the adjusted coefficient is multiplied by the mean value of PRODUCTION divided by the mean value of RDSTOCK. This is an estimate of the first derivative of productivity with respect to the stock of R&D taken at the means (see Appendix 2 for calculations). The procedure yields a figure of 58 per cent for the rate of return on R&D across all firms, with the use of the coefficient from the first productivity regression. Use of the separate RDSTOCK coefficients for the chemical products and electrical products sectors from the second regression produces rates of return of 64 per cent and 16 per cent respectively. These compare with a rate of return for capital (similarly computed) of 11 per cent.

There is no statistically significant relationship between R&D and own productivity for the firms in the sectors <u>not</u> broken out. Those firms have generally very low R&D stock-to-production ratios so this is not surprising. Since these firms contribute little to the value of the coefficient on R&D, it can be concluded that the R&D stock-to-production ratio of only the firms in the dummied sectors is more applicable for computing the rate of return. Therefore the productivity regression was re-estimated with just those firms. The results of this regression are shown in Table 4. The calculated rate of return on R&D using the coefficients from this regression and the R&D stock-to-production ratio for just the relatively R&D-intensive firms, is 24 per cent. This appears to be a much more

realistic figure, although the reader is cautioned that it applies only to the relatively R&D-intensive sectors.

The results thus provide evidence that the rate of return on R&D within a firm is well above that on capital in general, but this difference varies considerably among sectors. The estimates indicate that in order to equalize the rates of return on capital and R&D, the stock of R&D would have to rise by about 120 per cent. This increase in the R&D stock would lower the marginal rate of return on R&D to about 11 per cent, making it about equal to the estimated rate of return on capital. There is, however, reason to believe that the estimated return on capital is low (see Appendix 2). It has been estimated previously to be as high as 15 per cent. Given this higher figure for the rate of return on capital, the R&D stock would need to be augmented by about two-thirds to equalize the rates of return.

Table 3

Regression Results on Labour Productivity

Independent Variable	Productivity Regression 1		Productivity Regression 2	
	Coefficient	T-Value	Coefficient	T-Value
FO1 + FO2	0.037	(1.16)	0.029	(0.91)
ASSETS	0.095***	(5.87)	0.100***	(6.20)
PHW	-0.814***	(-38.74)	-0.822***	(-38.34)
MATERIALS	0.692***	(32.60)	0.697***	(31.96)
UTILIZE	0.224***	(5.67)	0.232***	(5.86)
RDSTOCK	0.020**	(2.19)	-	-
INTERCEPT	0.148	(0.79)	0.591**	(2.03)
SECTOR (1)	-0.068	(-1.29)	-0.200	(-0.61)
SECTOR (2)	-0.086	(-0.86)	-0.765	(-0.93)
SECTOR (3)	-0.087	(-1.55)	-0.674**	(-2.16)
SECTOR (7)	0.022	(0.43)	-0.699**	(-2.25)
SECTOR (9)	-0.039	(-0.50)	-0.459	(-1.09)
RDSTOCK (Other)	-	-	-0.012	(-0.63)
RDSTOCK (1)	-	-	0.008	(0.31)
RDSTOCK (2)	-	-	0.051	(0.88)
RDSTOCK (3)	-	-	0.044**	(1.94)
RDSTOCK (7)	-	-	0.055**	(2.35)
RDSTOCK (9)	-	-	0.032	(1.00)
R^2	0.960		0.963	
F-value	293.16		206.76	
Observations	146		146	
Degrees of Freedom	134		129	

** - Significant at the 95 per cent confidence level (two-tailed test).
*** - Significant at the 99 per cent confidence level (two-tailed test).

Table 4

Regression Results on Labour Productivity

Independent Variable	Productivity Coefficient	Regression 3 T-Value
FO1 + FO2	0.045	(1.11)
ASSETS	0.080***	(4.15)
PHW	-0.790***	(-24.29)
MATERIALS	0.689***	(24.02)
UTILIZE	0.282***	(4.31)
RDSTOCK	0.024**	(2.13)
INTERCEPT	-0.070	(-0.31)
SECTOR (1)	-0.006	(-0.08)
SECTOR (2)	-0.064	(-0.57)
SECTOR (3)	-0.056	(-0.76)
SECTOR (7)	0.082	(1.12)

R^2	0.943
F-value	164.69
Observations	110
Degrees of Freedom	99

** - Significant at the 95 per cent confidence level (two-tailed test).
*** - Significant at the 99 per cent confidence level (two-tailed test).

SUMMARY AND CONCLUSIONS

This examination of the R&D productivity nexus began with a theoretical analysis, which incorporated the primary components of the industrial organization view of R&D decision-making by business firms. The output of R&D is knowledge. This is always a public or semi-public good, since it is a socially non-rivalrous commodity whose value does not deteriorate through use. Industrial R&D is generally a semi-public good because the entry barriers that are created by R&D allow the firm some exclusive use of the fruits of its research. The entry barriers, whether they are supplemented by patent protection or not, allow the firm to internalize some, but never all of the benefits of R&D.

Other impediments to private sector funding of R&D include lags in the development process, leaks of knowledge to competitors, and the uncertainty inherent in R&D expenditure. Schumpeter maintained that the effects of these market imperfections could be mitigated by the firm's market power, other than that emanating from the R&D itself.

The analysis here shows that given only the weak assumptions that firms are profit-maximizers and that R&D is a semi-public good, the Schumpeterian hypothesis is correct. This implies that there is a trade-off between short-term and long-term efficiency objectives. More market power for firms can mean more static allocative inefficiency, but a better allocation of investment resources between knowledge and physical capital.

Empirical studies of the relationship between R&D and concentration or share variables have produced mixed results. Quantitative support for a generally positive relationship is weak at best. This, as noted before, does not disprove the Schumpeterian hypothesis. Market power is affected by how near, as well as how many, are a firm's competitors. It can, however, be inferred from this evidence that there would be no advantage to weakening competition policy, for the purpose of increasing R&D. Increased concentration does not have much effect on R&D.

The results of this study and others, do suggest, though not very strongly, that R&D does increase more than proportionately with size. However, previous US studies have generally found that this is only true up to a point; usually about $200 million in sales.[67] A test of this subsidiary hypothesis with Canadian data would not be very meaningful because of the paucity of observations above this size. In any case, competition policy is unlikely to affect a manufacturing firm's development up to this threshold.

Rather than suggesting an alteration of competition policy, the implication of the analysis and the accumulated empirical evidence is that R&D subsidization should be selective. This is because the factors that determine the social returns on subsidies vary. At the very least, support should be weighted toward those industries least able to appropriate the returns to R&D. This will occur in those areas in which risk is highest, where the knowledge is the least well protected by patents or other factors, and where development lags are the most discouraging. As well, because of its salient effects on competition and the possible existence of low-level economies of scale in the R&D process, subsidization should be weighted toward firms in the early development phases.

Some of the other parameter estimates also have specific policy implications. Evidence was found that federal grant programs generally cause Canadian firms to increase their privately financed expenditures on R&D by an amount equal to the grant. Conversely, foreign-owned firms do not respond with additional expenditures of their own to the grants. There is also very strong evidence that foreign ownership has a negative effect on R&D performance, other things being equal. This confirms the results of some earlier, less satisfying tests of this hypothesis. However, a sub-group of foreign-owned firms, those described as operating

under world product mandates or specialized missions, are not less R&D-intensive than similarly placed Canadian firms.

Another interesting result is the consistently positive relationship between payments for outside technology and R&D. This implies that R&D increases a firm's ability to obtain and absorb outside technology. In a small open economy like Canada's, a primary goal of industrial science and technology policy is to get maximum benefit from the international pool of technology. The finding that R&D and purchased technology are complementary suggests that technology transfer and diffusion are facilitated by indigenous research capability.

The production function analysis yielded two pieces of information that have important policy implications. First, there is no significant difference between the productivity of Canadian and foreign-owned firms when factor proportions and capacity utilization are accounted for. This suggests that the firm-specific advantages that are the raison d'être for transnational activity consist of advantages in financial capability or marketing activity rather than superior productive efficiency.

Finally, the marginal rate of return to the firm on R&D was found to be about 24 per cent in the more technology-intensive sectors. This is substantially above the rate of return to capital because firms discount returns to R&D owing to their relative riskiness. This implies that industrial R&D would have to be increased by about two-thirds over 1979 levels to equalize the rates of return to the alternative investments and maximize efficiency. Preliminary figures indicate that between 1979 and 1982, real industrial R&D rose by about 35 per cent, most likely because of increased government support.

The estimate of the underfunding of R&D given here is a very conservative one. The estimate does not account for the benefits of R&D that accrue to firms other than the performer. These external effects have been estimated in the US to be substantial. Further Canadian research should be directed toward producing comparable estimates for the Canadian economy. To do so will require the compiling of extensive additional data. Specific recommendations for further data development and research are contained in Appendix 2.

APPENDIX 1

The Minimum R&D Price

The derivation of the minimum R&D price is done using a method developed by Modigliani for finding the limit price that emanates from economies of scale.[68] To determine the minimum R&D price (P_r) only two equations are needed. Equation (i) specifies that when P_r is charged, profits are exactly equal to RD.r. This is the payment value of a perpetual annuity with a present value equal to RD, which is approximately equal to an "interest method" amortization of the cost of R&D over a long patent term. Equation (ii) gives the definition of the arc elasticity of demand (e_a) between X_c and X_r, which will be approximately equal to the point elasticity (e) at X_c.

(i) $\quad \text{Profit} = X_r.P_r - X_r.P_c = RD.r$

(ii) $\quad e_a = [(X_c - X_r)/X_c][P_c/(P_r - P_c)] \simeq e$

Because e is a constant, given the demand and cost functions, we have two equations in the two unknowns, X_r and P_r. For convenience, e is defined positively. In order to find P_r, both equations are solved for X_r as shown by (i)' and (ii)'.

(i)' $\quad X_r = RD.r/(P_r - P_c)$

(ii)' $\quad X_r = X_c - X_c(P_r - P_c)e/P_c$

Substitution of (i)' into (ii)' and some further rearrangement yields a quadratic in $(P_r - P_c)$. The result is equation (iii).

(iii) $\quad 0 = X_c(P_r - P_c) - X_c\,(e/P_c)\,(P_r - P_c)^2 - RD.r$

Solving (iii) for $(P_r - P_c)$ and moving P_c to the right side yields (iv), which gives zero, one or two real roots depending on the parameter values.

(iv)

$$P_r = P_c + P_c[X_c \pm (X_c^2 - 4X_c.e.RD.r/P_c)^{1/2}]/2X_c e$$

If there is no real root, the cost of R&D is greater than the monopoly profits. There will be a unique solution if monopoly profits are exactly equal to the cost of R&D. If the monopoly profits are more than sufficient to cover the cost of R&D there will be two roots, one above the monopoly price and one below. Obviously the minimum price necessary to support the expenditure on R&D will be the lower of the two. Thus equation (1) on page 13 is equation (iv) with the $\pm$ symbol replaced by a minus sign.

The Sylos Postulate

The possible responses of the inventing firm to imitative entry range from allowing the entrant all the room it wants by reducing output by whatever the imitator produces, to a predatory output increase designed to drive the competitor out of business. The first alternative is simple collusion. If this were realistic, there would be no point in analysing market structure; the simple assumption of classically monopolistic behaviour would suffice. The assumption that the initial firm holds output constant (hereafter called the Sylos Postulate following Modigliani)[69] is an intermediate case.

The rationale for the Sylos Postulate is quite simple. It is reasonable to assume that the inventor's plant capacity is sufficient to produce its pre-competition output at minimum short-run cost. Sunk costs of capital inputs that are given for some substantial period of time give the innovator an incentive to maintain output, in response to a fall in demand, at a level higher than that indicated by the junction of long-run marginal cost and marginal revenue.

Although the original price-output-capacity decision is based on long-run cost, once the capacity decision is taken, the firm reacts to any deviation from expected demand on the basis of variable costs alone. Because short-run marginal cost must be below long-run marginal cost (which includes capital costs), the intersection of marginal revenue with the former must be to the right of the intersection with the latter.

There are therefore, disincentives for the firm to reduce output. This makes the Sylos postulate seem quite reasonable, although it is obviously a caricature of actual performance. For a completely realistic behavioural specification, the parameters of the short-run cost curves would have to be known.

There are two obvious alternatives to the Sylos Postulate for this type of analysis: (a) the initial firm holds price constant, and (b) the firms form a cartel once entry has occured. The first alternative implies the complete naiveté of firms. This effectively assumes away the possibility of any kind of oligopolistic behaviour and is dismissed as unlikely for this reason. The alternative of collusion is cynical but implicitly assumes no cartel instability. Unless the stability of the cartel is accepted (as it could be in special cases), the Sylos Postulate is the most reasonable assumption of behaviour.

Derivation of the Limit Price for Homogeneous Products

The industry's demand curve is assumed to be linear as shown in (v). If a potential entrant expects the innovator to hold output constant when he introduces a homogeneous product, the entrant's demand will be given by (vi); where X_E refers to the entrant's output and X_L to the entrant's expectation of the output of the innovator, which is merely the innovator's output at the time of entry.

(v) $P = a_0 - a_1X$
(vi) $P_e = a_0 - a_1(X_E + X_L)$

It is necessary to find the quantity X_L that will just allow an imitator to enter and earn zero profits net of the R&D cost. This occurs when the entrant's maximum profits equal R&D. Equation (vii) gives the condition of the entrant's profit maximization. Marginal cost (P_C) is set equal to the expression for marginal revenue that can be derived from equation (vi). Equation (viii) states that the entrant's profits equal the cost of R&D.

(vii) $P_C = a_0 - 2a_1 X_E - a_1X_L$
(viii) $(P_E - P_C)X_E = RD.r$

Subtraction of (vii) from (vi) gives:

(ix) $(P_E - P_C) = a_1X_E$

substitution of (ix) into (viii) produces:

(x) $a_1X_E^2 = RD.r$

Solving (vii) and (x) for X_E and substituting the rearranged (vii) into the rearranged (x) gives an equation with X_L and constant parameters.

(xi) $(a_0 - P_C - a_1X_L)/2a_1 = (RD.r/a_1)^{1/2}$

This is solved for X_L to produce (xii):

(xii) $X_L = [a_0 - P_C - (4RD.r.a_1)^{1/2}]/a_1$

Equation (xii) shows the quantity that the initial firm has to produce to keep out an imitator. This is the limit quantity. To find the limit price, (xii) is merely substituted into (v), which is the industry demand and the demand the innovator initially faces.

(xiii) $P = P_C + (4RD.r.a_1)^{1/2}$

Thus (xiii) shows the limit price given by (2) on page 15.

Market Growth

If the market is growing as shown by equation (3) on page 15, the limit price will have to decrease over time if the ordinary Sylos Postulate is maintained. To derive equation (4), it is assumed that potential imitators take into account the market growth, so that equation (xii) is replaced by (xvi).

(xvi)
$${}_tX_L = [a_0(1 + g)^t - P_C - (4RD.r.a_1)^{1/2}]/a_1$$

However, the initial firm faces the current demand ($t = 0$), which is the same as the demand in the no-growth case. Thus (xvi) is plugged into (v) to give (xvii), which is the same as (4) if the subscript L is added to P.

(xvii)
$$P = P_C + (4RD.r.a_1)^{1/2} + a_0[1 - (1 + g)^t]$$

With the market taken as growing as shown by equation (3), the change in the hypothetical competitive equilibrium price (P_C) will be given by (xviii).

(xviii) ${}_tP_C = (1 + g)^t a_0 - a_1 {}_tX_C$

Solving this for ${}_tX_C$ gives the competitive equilibrium output.

(xix) ${}_tX_C = [a_0(1 + g)^t - {}_tP_C]/a_1$

Subtracting the competitive output of the static case from (xix) gives (xx), the second term in (5).

(xx) ${}_tX_C - {}_oX_C = [a_0(1 + g)^t - a_0]/a_1$

To find the limit price when there is market growth that is taken into account by the innovator as well as the imitator, (v) must be replaced by (3) before substitution of (xvi) is made. The growth effects then cancel out and the limit price is the same as for the no-growth case given by (2).

Limit Price Under Conditions of Product Heterogeneity

When products are heterogeneous, a firm's quantity demanded will be a function of its competitor's price as well as its own. As long as the relationship is monotonic in all variables, the equation can be inverted so that price is given as a function of quantity and the competitor's quantity. This is the way the relationships are shown in equations (6) and (6a) on page 18. Writing the demand relationships this way is merely a matter of convenience that makes the following manipulations simpler.

The limit price, when the entrant expects the innovator to hold output constant, is obtained in the same way as the quantity-held-constant limit price for a homogeneous product. As before we use the profit-equals-R&D cost and profit-maximization conditions to produce (xxi), which is a slightly rearranged version of (x) (except that b's instead of a's are now being used for the demand parameters). Equation (xxii) shows the marginal revenue of the entrant.

(xxi) $X_E = (RD.r/b_1)^{1/2}$
(xxii) $P_C = b_0 - 2b_1X_E - b_1b_2X$

By solving (xxii) for X_E and plugging the result into equation (xxi), (xxiii) is obtained. Rearrangement of (xxiii) yields (xxiv).

(xxiii) $(b_0 - b_1b_2X - P_C)/2b_1 = (RD.r/b_1)^{1/2}$
(xxiv) $X = [b_0 - P_C - (4RD.r.b_1)^{1/2}]/b_1b_2$

Equation (xxiv) gives the quantity of the innovator that will keep a potential entrant's profits to zero. This is substituted into the innovator's demand curve (from which the X_E term is dropped) to give (7a) on page 19.

The logic behind dropping the X_E term is that the innovator's quantity is set before there is any entry. If this is the case, the limit price will be determined by the entry-preventing output (given by the entrant's demand conditions without reference to the competitor's actions). To get the price-held-constant limit price, instead of dropping the X_E term, the right side of (xxi) is substituted for X_E. This yields equation (7).

APPENDIX 2

Data Sources

Four different confidential Statistics Canada data sources were used in this study. The information on R&D, government R&D grants and contracts, and technology payments and receipts came from the Science Statistics database. The variable ASSETS was taken from the CALURA (Corporate and Labour Unions Returns Act) database. The three-digit level variables were obtained from a special file constructed by the Multinational Enterprises Section of Statistics Canada in cooperation with the Economic Council of Canada. All the other variables are from the MAPID (Census of Manufacturers and Primary Industries) database.

The main problem in achieving a database with all the variables desired for this analysis was the fact that MAPID is collected on an establishment basis while CALURA and the Science Statistics are corporate-level databases.[70]

Although these cannot always be matched exactly at their respective reporting levels, they can be linked at the "enterprise" level. An enterprise is defined as the aggregation of all the reporting units held under common ownership. Thus the records for each of the databases were added on enterprise numbers.

The next step was to break out the MAPID data into "unconsolidated enterprises" (called firms). These consist of all the establishments of an enterprise that are assigned to a particular three-digit industry. The problem was then to assign the firms their proper proportion of the R&D of the enterprise.

This was done in a two-step procedure. First, assignments were made for those enterprises that either operate entirely in one three-digit industry or for which a survey done by the Science Statistics Centre (in 1980) gave us a product category breakdown of R&D expenditures. This enabled the proportioning of over 50 per cent of the R&D.

The rest of the R&D was then allocated on the basis of a double weighting procedure. The weights were the fraction of the output of the enterprise that a particular firm accounted for and the ratio of the firm's industry average R&D intensity (computed with the data already allocated) to the sum of the R&D intensities of the industries in which the enterprise operated.

Allocation of the other Science Statistics variables was done in the same fashion using industry intensities of the relevant variables.

There were 353 reporting units in the Science Statistics database in the 24 industries studied. These accounted for 88 per cent of all the manufacturing R&D done in Canada in 1979. The reporting unit for the Science Statistics is usually a corporation. However, divisions of corporations sometimes report separately. When collected into enterprises these 353 records were reduced to 252. The 252 enterprises were matched to the MAPID data, then broken out (as explained above) into 447 firms.

In order to estimate a production function, a figure for the capital stock is needed. For this reason the assets data from CALURA was used, since MAPID does not provide any kind of figure for capital stock. An alternative would have been to use the estimates from another Statistics Canada survey that gives numbers for net investment in physical capital. From this establishment-based survey an estimate of capital stock could be computed. This alternative would almost certainly have been better because the incompatibility of CALURA and MAPID is not entirely reconcilable. However, this data was not available for a sufficient number of the R&D performers.

Thus the CALURA records were summed to make corporate-level unconsolidated enterprises consisting of all the corporations in a given industry under common ownership. These were then linked directly to the MAPID firms. This meant a substantial reduction in the number of observations (236) since enterprises do not always have a corporation classified to every one of the industries in which they operate. There is undoubtedly considerable noise in the assets figure then,

because the corporate three-digit enterprise may include establishments not classified to that industry in MAPID or may alternatively be lacking establishments that are classified to that industry in MAPID. Moreover, the assets figure will probably be generally overestimated, because the corporate books include assets which are used for distributional rather than manufacturing activity. The implications of this are discussed below.

For the estimation of the production function, the data set was further reduced to 146 observations, by the elimination of firms for which R&D data back to 1972 was not available.

Lag Structure and Depreciation of the R&D Stock

Other lag rates tried were half a year, two years and three years. Each of these were tried with depreciation rates of 10, 12.5 and 15 per cent. The difference in fit between any two of the lag-depreciation combinations was insignificant.

Rate of Return Calculations

The calculation for the rate of return on a factor, given Cobb-Douglas technology, is $b_i(Y/X_i)$, where Y is output, X_i is the i^{th} factor and b_i is its ceteris-paribus coefficient. As explained in the text the presence of UTILIZE requires the adjustment of the estimated coefficients by multiplication by $1/(1 - b_4)$. These calculations are shown below, where RDSTOCK and OUTPUT are in millions of dollars.

Productivity Regression 1
(dollar figures in millions)

RDSTOCK = \$1,347
PRODUCTION = \$30,792
b_3 = .020
b_4 = .22

Return = .02/(1 - .224) X 30,792/1,347
= .58

Productivity Regression 2

Chemical Sector:

RDSTOCK(7) = \$316
PRODUCTION = \$3,617
b_3 = .055 - .012 = .043
b_4 =.23

Return = .043/(1 - .23) X 3,617/316
= .64

Electrical Sector:

RDSTOCK(3) = \$ 809
PRODUCTION = \$ 3,084
b_3 = .044 - .012 = .032
b_4 = .23

Return = .032/(1 - .23) X 3,084/809
= .16

Productivity Regression 3

RDSTOCK = \$1,235
PRODUCTION = \$8,786
b_3 = .024
b_4 = .28

Return = .024/(1 - .28) X 8,786/1,235
= .24

Return on Capital
from Productivity Regression 1

ASSETS = \$34,670
PRODUCTION = \$30,792
b_1 = .095
b_4 = .22

Return = .095/(1 - .22) X 30,792/34,670
= .11

The multiplicative nature of the Cobb-Douglas function means that a general over- or underestimate of a factor (multiplication of the factor series through by a constant) will not affect the slope coefficients. The

effect will be buried in the intercept. However, the rate-of-return calculation will be affected. Thus the estimate of the rate of return on R&D could be either high or low, depending on the accuracy of the depreciation and lag rates chosen.

A solution to this problem would be a form of mixed estimation in which estimates of these factors were taken from another study designed specifically to produce them. Unfortunately, no study appropriate for this purpose exists, so that the alternative of choosing on the basis of fit was all that was available.

An Alternative Estimation of the Production Function

A transformation of the Cobb-Douglas production function allows estimation of the rate of return on R&D without the necessity of calculating an R&D stock.[71] Consider the production function shown below, where Y, K, L and S are value added, capital, labour and the stock of R&D respectively.

$$Y = AK^{u}L^{v}S^{w}$$

This can be expressed in instantaneous rate of change form. The lower case letters refer to the rates of change.

$$y = uk + vl + ws$$

The instantaneous rate change of the R&D stock (s), can be approximated by the annual net addition to R&D over the stock of R&D at the beginning of the year (RD/S).

Thus ws can be replaced by w(RD/S), which can then be replaced by r(RD/Y), where r is the rate of return on R&D.

Proof: $w(RD/S) = w(Y/S)(RD/Y)$
$= r(RD/Y)$; by definition of r.

If zero depreciation is assumed, RD will be simply the annual expenditure on R&D. Thus only R&D expenditure data is needed to estimate the equation since the stock of R&D has been removed by the substitution shown above.

One disadvantage of this specification is that the rate of change relationship will be generally noisier since less of the relevant information is used. This has to be weighed against the elimination of the bias (unknown in degree or direction) emanating from the necessity of estimating the stock of R&D. Though if a sufficient number of observations are available, the relative noisiness of the rate of change form will not be a problem.

There is one other major difference between the forms. The rate of return on R&D is forced to be a constant across firms in the alternative specification. In practice this will probably be a highly restrictive assumption. It can, however, be tested without using the rate of change form.

In the standard log estimate, the R&D term is $w\log_e S$. This can be replaced by $r(S/Y)\log_e S$. If this is estimated, the rate of return will be forced to equality for all firms. Furthermore, this hypothesis can be nested with the alternative that the rate of return is unequal across firms. This requires the two terms shown below.

$$h.r(S/Y)\log_e S + w(1 - h)\log_e S$$

If $h = 1$ the equality hypothesis is correct, and if $h = 0$ there is no tendency for the returns to be equal. The estimate will include two R&D terms:

$$(S/Y)\log_e S \text{ and } \log_e S.$$

If we call the estimated coefficients c_1 and c_2, then

$$c_1 = h.r$$
$$= h.w(Y/S)$$

and

$$c_2 = w(1 - h)$$

If Y and S are replaced with their mean values, there will be two equations in two unknowns once values for c_1 and c_2 are obtained. Hence, values for w,r and h will be produced.

When this regression was tried, the value of c_2 was significant and the value of c_1 insignificant. The values of w and r were close to what were obtained with the straight Cobb-Douglas specification, and the value of h was close to zero. As well, when the equation was estimated with only the $(S/Y)\log_e S$ term, the coefficient was insignificant. It was therefore concluded that the rate of return at the margin varies considerably among firms in Canada because of differing factor intensities, so that the alternative specification was not beneficial.

Recommendations for Data Development and Further Research

As mentioned a number of times in this study, the availability of data on many important economic variables is unsatisfactory. This is particularly frustrating because many of the most serious drawbacks of the existing data sets could be corrected with only a small amount of additional work by Statistics Canada and the reporting units. What follows are some practical suggestions for the augmentation of current data sets that could produce much improved estimates of key policy parameters. At the end of this section some comments will also be made on the best approach to estimating the social returns to R&D, given the constraints of the data.

One obvious and serious gap in our knowledge of R&D is the inflation rate for R&D expenditures. This should be corrected as soon as possible. R&D performance is concentrated in a relatively small number of industries, therefore not much dissagregation would be required in the R&D deflators. Six different indices should cover industrial R&D reasonably well: aircraft and aircraft parts, electrical products, non-electrical machinery, chemical products, all manufacturing and all industrial R&D.

Another problem with the R&D statistics is industry breakdown. The Science Statistics are collected at the corporate level. Corporations are legal fictions - their accounting data often agglomerates the results of operations in a large number of different industries. To allocate R&D expenditures properly, the Science Statistics survey should include an additional page for manufacturing corporations that asks the respondent to break down R&D by three-digit SIC. On the same page the respondents should be asked for a breakdown between product and process R&D.

Since most manufacturing R&D is done in only 24 out of 89 three-digit industries, the page should present a matrix with 25 rows, corresponding to the industries (and including an "other" category) and two columns headed product and process.

The product and process terms will, of course, have to be defined for the respondent, who will be asked to assign percentages of the corporation's R&D among the 50 cells in the matrix. For most corporations this will not be a particularly onerous task since its R&D can be assigned to a small number of cells, with the rest left blank.

The above suggestions for improving the R&D data are similar to those made by H.H. Postner in a recent discussion paper for the Economic Council.[72] Postner, however, concerned himself with recommendations for the long-term improvement of data collection. In contrast, the recommendations given here focus on minor adjustments to current practices that could produce improved estimates within the next couple of years. These two objectives are not incompatible. In fact, they should be considered complementary. Any overhaul of data collection practices must include ways to fine-tune the existing system, because public policy always needs to be made yesterday.

There are also serious gaps in the data on production from the Census of Manufacturers and Primary Industries (MAPID). This is generally a very good database. Time-series analysis is easy with MAPID (as it is with the Science Statistics), because the records carry inter-temporally consistent identification numbers. More importantly, MAPID is establishment-based. This supports the concept of the firm as an unconsolidated enterprise, which provides a much better approximation of a "line of business" than

the corporate fiction on which tax data is based. However, two variables that are crucial for productivity analysis are not available in the MAPID database. These are capacity utilization and physical capital stock.

There is no survey measure of capacity utilization anywhere in Canada. Both Statistics Canada and the Department of Regional Industrial Expansion produce indirect estimates of capacity utilization at the two-digit level. However, the crude methods used to produce these estimates make them unsatisfactory at the macro level and useless for microeconomic analysis. This problem could be solved with the addition of just one query to the MAPID questionnaire.

The net book value of physical capital stock could also be obtained with one additional question. This would be very useful for improving micro-level analysis of productivity, since the investment numbers produced by the construction division are geared toward producing aggregates, and hence are not comprehensive at the establishment level.

Probably the most important missing link in our knowledge of R&D in Canada is an estimate of the social rate of return to R&D. In order to make a good estimate of this, a "technology flows" matrix is needed. Construction of this matrix will require the improvements in the Science Statistics suggested above as well as the use of patents statistics. The PATDAT database contains information on the use, origin and type (i.e., product or process) of all Canadian patents since 1978.

What is actually required for estimating the rate of return to R&D is a tracing of R&D flows themselves. Either production flows or patents flows can provide a crude approximation of R&D flows. However, both of these proxies are badly flawed. A much better approximation of R&D flows could be accomplished by making use of both the patents and production flows.

The integration of this information would involve the construction of indices that are weighted averages of the production and patents flows indices. The weight put on an industry's patents index would be the proportion of R&D that served to produce the patented innovations. The remaining R&D would then be assigned on the basis of the production flows. This method would make use of the important information contained in PATDAT, while accounting for the fact that patents are of widely different importance in different industries.

The weights would require some new survey data, but not a whole new survey. Firms could be asked on the Science Statistics questionnaire to give the proportion of their R&D that went into the innovations that they patented. An even better approach would be to include in the file for each patent an estimate of the research and development costs that went into it.

This is useful research that is well worth the cost of the extra data development it would require.

NOTES

1. Donald G. McFetridge and Jacek P. Warda, Canadian R&D Incentives: Their Adequacy and Impact (Toronto: Canadian Tax Foundation, 1983), 63.

2. Ibid., 76.

3. Ibid., 83.

4. Research and Development Tax Policies, April 1983 Federal Budget, (Ottawa: Department of Finance, 1983), 27.

5. Jora R. Minasian, "Research and Development, Production Functions, and Rates of Return," American Economic Review 59 (1969): 80-86, also references in notes 58 and 60.

6. The first of these studies was Nestor E. Terleckyi, Effects of R&D on the Productivity Growth of Industries: An Exploratory Study, National Planning Association, 1974, 44. See also references in notes 10, 55 and 56.

7. M. Kamien and N. Schwartz, "Market Structure and Innovation," Journal of Economic Literature 13 (1975): 32.

8. R.E. Caves et al., Competition in the Open Economy: A Model Applied to Canada (Cambridge, Mass.: Harvard University Press, 1980), 193; and Rosalyn Frankl, "A Cross-Section Analysis of Research and Development Intensity in Canadian Industries with Particular Reference to Foreign Control" (Ottawa: Department of Industry, Trade and Commerce, 1979), 40.

9. H.H. Postner and L. Wesa, Canadian Productivity and Growth Analysis (Ottawa: The Economic Council of Canada, 1983), 29-36.

10. F.M. Scherer, "Inter Industry Technology Flows in the United States," Research Policy 11 (1982): 227-245.

11. The cost of the R&D could be added to the LAC schedule in the form of an annuity over the life of the patent with a present value of the cost of the R&D. However, the solution would remain the same because marginal cost would still equal the net of R&D average cost in the efficient scale range.

12. J.S. Bain, Barriers to New Competition (Cambridge, Mass.: Harvard University Press, 1956), 3.

13. Ibid., 4.

14. A. Koutsoyiannis, Modern Microeconomics, 2nd edition (London: Macmillan Press Ltd., 1979), 551.

15. With economies of scale, but no long-run diminishing returns, the shares are indeterminate within a range. The rationality assumption implies only that firms will enter at least at minimum efficient scale.

16. P. Sylos-Labini, Oligopoly and Technical Progress (Cambridge, Mass.: Harvard University Press, 1957).

17. G.J. Stigler, The Theory of Price, 3rd edition (New York: The Macmillan Co., 1966), 344.

18. B. Branch, "Research and Development Activity and Profitability: A distributed Lag Analysis," Journal of Political Economy 82 (1974): 999-1011.

19. Scherer, op. cit., 227-245.

20. D.P. DeMelto et al., Innovation and Technological Change in Five Canadian Industries (Ottawa: Economic Council of Canada, 1980), 26.

21. A uni-directional barrier to entering the recipient country would suffice. Only if there were no barriers would the exporter have no ability to differentiate the markets.

22. It is assumed that the foreign producer has greater market power since it is able, by assumption, to penetrate the domestic market despite trade barriers.

23. E. Mansfield, "How Economists View R&D," Research Management, 25 (1982): 24.

24. The variance of the average rate of return of a large number of projects will approach zero unless projects' returns are generally positively correlated with each other.

25. M. Kamien and N. Schwartz, Market Structure and Innovation (Cambridge: Cambridge University Press, 1982), 22.

26. Kamien and Schwartz, op. cit., 9.

27. Ibid., 32.

28. J.D. Howe and D.G. McFetridge, "The Determinants of R&D Expenditures," Canadian Journal of Economics 9 (1976): 68.

29. Caves et al., op. cit., 166.

30. Ibid., 24.

31. Caves et al., op. cit., 194.

32. Howe and McFetridge, op. cit., 68.

33. K.S. Palda and B. Pazderka, Approaches to an International Comparison of Canada's R&D Expenditures (Ottawa: Economic Council of Canada, 1982), 38.

34. F.M. Scherer, Industrial Market Structure and Economic Performance (Boston: Houghton Mifflin Company, 1980), 434.

35. Caves et al., op. cit., 175.

36. Kamien and Schwartz, "Market Structure and Innovation," 7.

37. P. Telesio, Technology Licensing and Multinational Enterprises (New York: Praeger, 1980), 110.

38. Caves et al., op. cit., 22.

39. It is assumed that a "before R&D" figure is used for profits. Current R&D would be an expense that reduces current profits otherwise. This would not be a problem of simultaneity, but rather identity.

40. These are: Caves et al., op. cit., 194; Howe and McFetridge op. cit., 67; and G.K. Hewitt, R&D in Selected Canadian Industries (Ottawa: Department of Industry, Trade and Commerce, 1981), 24.

41. Howe and McFetridge, op. cit., 66.

42. K.S. Palda and B. Pazderka, "Foreign Ownership and R&D Intensity: Analysis of International Data" (Kingston, Ont.: Queens University, School of Business, 1982), 1.

43. Caves et al., op. cit., 193; and R. Frankl, "A Cross Section Analysis of Research and Development Intensity in Canadian Industries with Particular Reference to Foreign Control" (Ottawa: Department of Industry, Trade and Commerce, 1979), 40.

44. Caves et al., op. cit., 186.

45. Howe and McFetridge op. cit., 67.

46. Kamien and Schwartz, op. cit., 25.

47. Palda and Pazderka, op. cit., 37.

48. Caves et al. op. cit., 186.

49. Caves et al. op. cit., 184-185.

50. Ibid., 185.

51. Howe and McFetridge, op. cit., 66.

52. Theoretically the relationship between PTRD and TECHPAY is simultaneous. However, it is not possible to identify separate regressions with the available variables. All that can be said is that the two variables are positively related, holding other factors constant.

53. N. Terleckyj, Effects of R&D on the Productivity Growth of Industries: An

Exploratory Study (Washington: National Planning Association, 1974), 44.

54. Scherer, op. cit., 227-245.

55. Ibid., 244.

56. E. Mansfield et al., "Social and Private Rates of Return from Industrial Innovation," Quarterly Journal of Economics 91 (1977): 221.

57. H.H. Postner, and L. Wesa, Canadian Productivity Growth Analysis (Ottawa: Economic Council of Canada, 1983), 29-36.

58. S. Globerman, "The Empirical Relationship Between R&D and Industrial Growth in Canada," Applied Economics 4 (1972): 186. After this study was completed, the author became aware of a paper presented by Lorne Switzer at the 1984 CEA conference, which did find a significant positive relationship between R&D and own productivity at the two-digit industry level.

59. See K. Clark and Z. Griliches, Productivity Growth and R&D at the Business Level: Results from the P.I.M.S. Database, (Washington: National Bureau of Economic Research, 1982), 31; J. Minasian, "Research and Development, Production Functions, and Rates of Return," American Economic Review 59 (1969): 80-86; and A. Link, "Inter-Firm Technology Flows and Productivity Growth," Economic Letters 11 (1983): 179-185.

60. Ibid.

61. J. Bernstein and M.I. Nadiri, "Rates of Return on Physical and R&D Capital and Structure of the Production Process: Cross Section and Time Series Evidence" (Ottawa: Carleton University, 1984), 2.

62. Ibid., 20.

63. Ibid.

64. All dollar variables were adjusted by the appropriate deflators to put them all in 1979 dollars before any manipulations were performed. The GNE deflator was used on R&D since no price index for R&D specifically is available.

65. In the second regression, RDSTOCK (other) is the partial regression coefficient on RDSTOCK. This can be interpreted as the slope w.r.RDSTOCK for the industries not broken out. For an industry (i) that is dummied RDSTOCK (Other) + RDSTOCK (i) gives the slope.

66. Bernstein and Nadiri, op. cit., 17.

67. Kamien and Schwartz, Market Structure and Innovation, 75-84.

68. F. Modigliani, "New Developments on the Oligopoly Front," Journal of Political Economy 66 (1958): 215-232.

69. Ibid., 215.

70. An establishment is defined as the smallest unit that is a separate operating entity.

71. Clark and Griliches, op. cit., 15.

72. H.H. Postner, "Statistical Problems of Relating Research and Development Data to Productivity Data" (Ottawa: Economic Council of Canada, 1983), 14.